Arno Bohm · Piotr Kielanowski
G. Bruce Mainland

Quantum Physics

States, Observables and Their Time Evolution

Springer

Quantum Physics

Arno Bohm • Piotr Kielanowski •
G. Bruce Mainland

Quantum Physics

States, Observables and Their Time Evolution

Springer

Arno Bohm
Department of Physics
University of Texas
Austin
TX, USA

Piotr Kielanowski
Department of Physics
CINVESTAV
Mexico City
Mexico

G. Bruce Mainland
Department of Physics
Ohio State University
Columbus
OH, USA

ISBN 978-94-024-1758-6 ISBN 978-94-024-1760-9 (eBook)
https://doi.org/10.1007/978-94-024-1760-9

© Springer Nature B.V. 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature B.V.
The registered company address is: Van Godewijckstraat 30, 3311 GX Dordrecht, The Netherlands

Preface

> "Quantum Mechanics" is the description of the behavior of matter and light in all its details and, in particular, of the happenings on an atomic scale. Things on a very small scale behave like nothing you have any direct experience about. They do not behave like waves, they do not behave like particles, they do not behave like clouds, or billiard balls, or weights on springs, or like anything you have ever seen.—*Richard Feynman*

This book has been written to serve as a text for an introductory graduate course; much of the material can also be used for an advanced, undergraduate course. Its precursor[1] has been used many times in a 1-year, introductory, graduate course.

Throughout the text either theoretical developments have been motivated by experimental observations or theory has been used to explain experimental data. The authors have made a significant effort to make the material more easily accessible by (a) providing systematic explanations, (b) presenting logical step-by-step derivations, (c) imbedding solved examples in the text at appropriate places to clarify the ongoing discussion, (d) summarizing key ideas at the end of each chapter, and (e) providing an extensive set of problems at the end of each chapter.

Some chapters may be appropriate for courses in quantum chemistry since many physical applications in this book have been chosen from molecular spectra and molecular structure,[2] which serves as a fertile ground for examples of non-relativistic quantum physics.

The sections "Precession of a Spinning Particle in a Magnetic Field: The Interpretation of the Schrödinger and Heisenberg Pictures" and "Magnetic Resonance" can be used in a special topics course as an introduction to nuclear magnetic resonance.

Complicated quantum systems can often be more profitably analyzed, not in terms of their constituents, but instead in terms of more general substructures: specifically, molecules can be analyzed in terms of rotations, oscillations, and

---

[1] A. Bohm, *Quantum Mechanics: Foundations and Applications*, Springer-Verlag, New York, 2nd Edition (1986), 3rd Edition (1993), soft-cover printing (2001).

[2] G. Herzberg, *Molecular Spectra and Molecular Structure*, Van Nostrand, New York, 1939–1966.

single-electron excitations,[3] and nuclei can be described in terms of their collective motions.[4] It is remarkable that this way of understanding can also be extended to relativistic systems where the hadron spectra can be analyzed in terms of relativistic rotators and oscillators.[5] This analysis of quantum physical systems in terms of collective motions rather than constituents is also used here to analyze molecular spectra and molecular structure in terms of rotational and vibrational motions.

The objective of this book is to present quantum mechanics in its general form by stressing the operator approach. A major new development in physics usually necessitates a corresponding development in mathematics. For example, differential and integral calculus were developed for classical mechanics, and no one today would teach an advanced course without using this mathematical language. Although the mathematics of linear, scalar-product spaces and linear operators were created and developed to fulfill the needs of quantum physics, quantum mechanics is still often taught without using its mathematical language. By restricting much of a course on quantum mechanics to differential equations and a discussion of their solutions, students may initially find the material easier to grasp because the mathematics is familiar. However, many quantum concepts are difficult to present in this narrow mathematical language that emphasizes only one of the many complementary aspects of quantum physics. There is much more to quantum mechanics than the overemphasized, wave-particle dualism presented in terms of differential equations, and there is no principle that a priori places position and momentum at the forefront. Every observable needs to receive the emphasis that is appropriate for the particular situation being considered.

The mathematics of quantum mechanics, which involves linear, scalar-product spaces and algebras of linear operators, is discussed in Appendix. Rather than treating the mathematics abstractly, each operation in a general, linear space is motivated by first examining the corresponding operation in the familiar three-dimensional vector space. Also, when the properties of scalar-product spaces are discussed, each property is first shown to exist both for the scalar product in three-dimensional vector space and for the scalar product expressed as an integral. The emphasis is on providing an introduction to the mathematics required to perform quantum calculations, not on providing mathematical justification (proofs). This elementary mathematical tutorial has been written for the reader who has no prior knowledge of the general mathematical structure of quantum mechanics. Thus the reader who has some familiarity with the mathematics of linear operators in linear, scalar-product spaces can skip the Appendix. If the reader then finds some aspect of the mathematics in Chap. 1 or in later chapters unfamiliar, the Appendix can be used as a reference.

The discussion of quantum mechanics begins with some basic postulates of quantum mechanics that are formulated and made plausible by using the example of

---

[3] Ibid.

[4] A. Bohr, B. R. Mottelson, and J. Rainwater, 1975 Nobel Prize in Physics.

[5] A. Bohm, Y. Neeman, A. O. Barut et al. *Dynamical Groups and Spectrum Generating Algebras*, World Scientific Publishing Co., 1988.

the harmonic oscillator realized by the diatomic molecule. Further basic postulates are introduced in later chapters when the scope of the theory is extended. These basic postulates are not mathematical axioms from which all predictions of the theory can be derived. Such an axiomatic approach does not appear to be possible in physics. Instead, the basic postulates are a concise way of expressing the essence of many experimental results and the successes of various theoretical ideas. In Chap. 1 representations of the algebra of the harmonic oscillator are first determined. The interpretation of experimental data from an energy loss experiment is used to motivate the introduction of the statistical operator—with matrix elements that form the "density matrix"—and to formulate the relationship between average values measured in an experiment and expectation values calculated theoretically. Radiative transitions between harmonic oscillator energy levels and the Einstein coefficients are discussed.

Representations of the algebra of angular momentum are calculated in Chap. 2. The algebra of angular momentum is enlarged by adding the position operator so that the algebra can be used to describe rigid and non-rigid rotating molecules. Theoretical predictions are compared with the experimental spectra of rotating diatomic molecules.

The combination of quantum physical systems using direct-product spaces is discussed in Chap. 3. The theory is used to describe a vibrating rotator, and the theoretical predictions are then compared with data for a vibrating, rotating diatomic molecule. The addition of angular momentum (Clebsch-Gordan coefficients) is discussed. Tensor operators are introduced so that the Wigner-Eckart theorem can be used to relate various experimental data.

The formalism of first- and second-order, non-degenerate perturbation theory and first-order, degenerate perturbation theory is derived in Chap. 4. The basic ideas associated with stationary perturbation are motivated by examining a rotator in a uniform magnetic field. Perturbation theory is used to explain the Stark effect.

Time development is described in Chap. 5 using either the time-dependent Schrödinger equation or Heisenberg's equation of motion. The Schrödinger picture, Heisenberg picture, and interaction picture are discussed. The precession of a spinning particle in a magnetic field is described in both the Schrödinger and Heisenberg pictures to help clarify the relationship between the two pictures. Magnetic resonance is discussed in the Schrödinger picture. The Gibb's distribution and a magnetic resonance experiment are discussed.

Since a discussion of the experimental and theoretical developments that presaged quantum mechanics is necessarily brief, in this book prior knowledge of classical mechanics, some electromagnetic theory, and some atomic physics is required. A basic knowledge of differential and integral calculus is assumed. Some familiarity with matrices, vector algebra, and linear spaces would be helpful.

Austin, TX, USA	Arno Bohm
Mexico City, Mexico	Piotr Kielanowski
Columbus, OH, USA	G. Bruce Mainland

Contents

Chapter 1
Quantum Harmonic Oscillator

1.1 The Gate to Quantum Mechanics

The story of quantum physics began in 1900 when Max Planck discovered by the thermodynamical methods the *improvement* of the Wien's law of energy distribution for blackbody radiation and then formulated the microscopic derivation of his equation in terms of oscillators within the cavity of a blackbody. Planck assumed that the energy of an oscillator of a given frequency ν has to be an integer multiple of an *energy element* $h\nu$

$$E = nh\nu = n\hbar\omega\,. \tag{1.1.1}$$

The angular frequency $\omega = 2\pi\nu$ and $\hbar = h/2\pi$ where the constant h is Planck's constant and has the experimental value

$$h = 6.626 \times 10^{-34}\text{Js}, \quad \hbar = 1.055 \times 10^{-34}\text{Js}\,. \tag{1.1.2}$$

Planck's assumption of discrete energies ran contrary to the ideas of classical physics where energy is continuous. He made it as "an act of desperation" in order to derive an energy distribution for blackbody radiation which agreed with experiments. Planck's assumption related to the oscillators lining the cavity of a blackbody radiator, but it is not far fetched to conclude that the energy emitted by such oscillators will also occur in amounts $h\nu$.

Still, 5 years elapsed before Planck's constant h was used again when Einstein (1905) explained the photoelectric effect by treating light as quanta with energy $E = h\nu$. If electromagnetic radiation consists of particles, now called photons, then these particles should possess momentum. Maxwell's classical wave theory of electromagnetic radiation leads to the result

$$p = \frac{E}{c}\,, \tag{1.1.3}$$

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9_1

where p and E are, respectively, the momentum and energy content in a given volume of a radiation field and c is the speed of light. Equation (1.1.3) is also, of course, just the usual relativistic equation for energy, $E^2 = p^2c^2 + m^2c^4$, when the particle's rest mass m is zero. If a photon, as Planck and Einstein suggested, has an energy given by (1.1.1), then from (1.1.3) its momentum is given by

$$p = \frac{h\nu}{c}. \tag{1.1.4}$$

Using the standard wave relationship $c = \lambda\nu$, which expresses the speed c of the wave in terms of its wavelength λ and frequency ν, (1.1.4) becomes

$$p = \frac{h}{\lambda}. \tag{1.1.5}$$

The "photon-electromagnetic wave" is an example of a physical system called an elementary particle. ("Particle" is an unfortunate choice for a name since it leads to an association with classical particles and the "photon-electromagnetic wave" is neither a classical particle nor a classical wave. It would have been better if a new name such as quanta had been coined.)

Since the seventeenth century two theories of the nature of light existed. Newton (1663) proposed that light existed as "corpuscles", or particles, and this came to be known as his corpuscular theory of light. Later Huygens (1678) proposed an alternate theory which considered light to be of a wave nature. At the time all physical facts known about light could be explained satisfactorily by either theory.

Two-slit interference of light, which could only be explained by the wave picture, was discovered by Thomas Young (1801). Newton's theory was unable to explain diffraction so the wave theory was generally accepted until the turn of the century. However, the photoelectric effect first observed by Heinrich Hertz (1887), and quantitatively established by P. Lenard (1902), could only be explained by viewing light as particles. Thus, two-slit interference and the photoelectric effect confront physicists with experimental results that cannot be simultaneously explained by either the wave picture or the particle picture alone. A theory is required that encompasses both the classical particle theory and the classical wave theory.

Later a similar dilemma arose concerning the wave and particle nature of the electron. From electrolysis experiments it was suggested that electric charge occurs only in discrete amounts. J.J. Thompson and J. Perrin (1897) showed that the phenomenon of cathode rays (electrons) could be explained if they are assumed to be a stream of particles, each having a definite mass and charge and obeying the laws of classical mechanics. This was understood as establishing the existence of the electron as a classical particle. The perception of the electron as a particle was further strengthened when Millikan (1909) measured the charge of a single electron and verified that charge in any region of space always comes in integer multiples of the elementary unit of change e. Thus by 1920 the electron was accepted as a particle because the particle picture described all known phenomena, whereas the wave picture could not explain the discrete nature of charge.

Physicists were then forced to question their acceptance of the electron as a particle when electron diffraction was discovered by Davisson and Germer and by G.P. Thompson (1927). These experiments could not be explained by the particle picture of electrons. However, the results could be explained by the wave picture if electrons with momentum $p = mv$ have a wavelength λ given by

$$\lambda = \frac{h}{p} . \tag{1.1.6}$$

The above relationship is identical in form to (1.1.5); however, (1.1.5) describes light while (1.1.6) describes electrons. Since $\lambda = h/p$ both for light and for electrons, it is reasonable to postulate that it is true for all objects. Defining the wave number k by the relation

$$k = \frac{2\pi}{\lambda} , \tag{1.1.7}$$

allows (1.1.6) to be written in the form

$$p = \frac{h}{\lambda} = \hbar k . \tag{1.1.8}$$

Historically, before the experimental discovery of electron diffraction, Louis de Broglie (pronounced de Broyle) (1924) conjectured, in analogy with the wave-particle duality of light, that particles possess wave properties and that their wavelength is given by (1.1.6). He saw the symmetry of nature and thought that the wave-particle duality of light should be matched by the wave-particle duality of electrons.

The Planck relation (1.1.1) and the de Broglie relation (1.1.8), namely

$$E = h\nu \quad \text{and} \quad p = \hbar k , \tag{1.1.9}$$

are justly considered the gateway to quantum theory.

In some experiments cathode rays exhibit both wave and particle properties: cathode rays, which consist of individual electrons, exhibit wave properties when they are scattered from a crystal surface, creating a diffraction pattern, or when they pass through two slits to create a two-slit interference pattern. But after passing through the slits, cathode rays exhibit particle properties when individual flashes are observed as single electrons strike a luminescent screen.

Furthermore some experimental results can be described equally well using either the particle or wave picture. For example, by examining the deflection of alpha particles off of thin metal foils, Rutherford concluded that almost all of the mass of an atom is concentrated in a small region. The formula that correctly predicts the amount of deflection for the alpha particles can be derived using either the particle picture or the wave picture. Also, the deflection of cathode rays in electromagnetic fields can be explained using either picture.

Both pictures have redundant structures, structures that cannot be observed in an experiment. When an electron is viewed as a particle, it apparently has a trajectory. But in attempting to determine the trajectory, for example by using light, the trajectory of the electron is changed when it is collides with a photon. The trajectory is unobservable since, in the process of observing it, the trajectory is changed. Similarly, when a beam of electrons is described as a wave, the matter density is treated as if it were homogeneous although the beam is comprised of individual electrons.

At the advent of quantum mechanics, light was thought to be a wave and electrons were thought to be particles only as a consequence of which experiments had been done by that time. If the photoelectric effect had been discovered before two-slit interference of light was observed, light would have been classified as a particle. Similarly, if electron diffraction had been discovered first, electrons would have been classified as waves. It may seem that since cathode rays and light can be described either as particles or as waves, the only function of a new theory is to reveal when to use the particle description, when to use the wave description, and when it does not matter. This, however, is not the case. Both for light and electrons, neither the particle theory nor the wave theory completely describes the observed behavior. A theory is needed that embraces the two. Quantum theory resolves this apparent paradox and shows that there is much more than just waves and particles. The question, "Are light and electrons waves or particles?" has been resolved, not by deciding it, but by showing that it is not the right question to ask.

Atomic dynamics began with the work of Bohr (1913). From the Rydberg formula it was known that the reciprocal of the wavelengths of light emitted and absorbed by hydrogen atoms is given by the difference of two terms, each of which depends on an integer. Consequently, discrete numbers (integers) play a role in atomic physics. To these integers Bohr assigned a discrete set of *states* that he believed were stationary orbits in which the laws of classical electrodynamics did not apply so that the electron would not radiate away its energy, causing the collapse of the atom. Thus the electron in the hydrogen atom could only be in one of a discrete set of states. Each atomic state of the electron has its own specific value of energy and orbital angular momentum and is called an eigenstate of energy and orbital angular momentum. ("Eigen" is the German word for "own," and an eigenvalue is the characteristic value or proper value.) An atomic state is different from a wave with a definite value of linear momentum $\mathbf{p}$ or a particle state with a definite value of position $\mathbf{x}$. The laws for observables in atomic states have a form that is very different from laws in classical physics. Therefore, a guiding principle is needed to conjecture these new relationships between atomic observables. This guiding principle is the Bohr correspondence principle.

The correspondence principle is one of the more fruitful conceptual devices employed in the development of a physical theory and played a major role in the discovery of quantum mechanics. The principle asserts that structures of the new, quantum theory have some correspondence to the structures of the old, classical theory. Relations between quantum observables can be found (conjectured) by using relations between the corresponding classical observables. Based on the observation that classical physics is correct for macroscopic objects, Bohr required that the new quantum theory makes predictions that agree with classical physics in the limiting cases of large masses or large orbits. From the Bohr model of the hydrogen atom the radii of the stationary orbits are proportional to the square of the quantum number n, so large radii correspond to large values of n. Therefore, the first part of the correspondence principle was stated in the following form:

> **Correspondence Principle, Quantum Numbers**: The predictions of quantum theory must correspond to the predictions of classical physics in the limit of large quantum numbers.

Classically it can be shown for certain systems that transitions are possible, not between any two states, but only between states that are related to each other in a specific manner. Rules specifying which state may transform (decay) into another are called selection rules. The second part of the correspondence principle concerns selection rules:

> **Correspondence Principle, Selection Rules**: A selection rule that is necessary to obtain the correspondence in the classical limit (large n) also applies in the quantum limit, (small n).

The correspondence principle will be used, both to find the observables for a quantum physical system and even to give a name to the system. Bohr's correspondence principle has become a guiding principle in the search for new theories in physics.

Atomic and molecular physics are conveniently described by a third picture that differs from both the particle and wave pictures. The atomic-molecular picture is best illustrated by considering electrons in atoms where neither the particle picture nor the wave picture is adequate. Such electrons are not detected with counters and do not undergo diffraction. Out of a plurality of states, electrons in atoms are in a third kind of state that is neither a wave state nor a particle state. This third type of state was of great importance historically in the discovery of quantum mechanics by Heisenberg (1924) that represents a significant advance over Bohr's semi-classical treatment of the hydrogen atom. Heisenberg's work provided a matrix formulation of quantum mechanics that is one form of quantum theory as it now exists.

The particle, wave, and atomic-molecular pictures do have features in common, and the primary purpose of this book is to discuss the underlying structure that these three approaches share. Remaining at the present level, where only external appearances are observed, it will not be possible to determine the universal structure of quantum mechanics. Instead one must delve more deeply into the mathematics of quantum mechanics and rid oneself of the notion that all observable quantities are represented by numbers. At this new, deeper level the confusion and apparent contradictions vanish. Reaching this deeper level is especially difficult because, when initially studying quantum mechanics, an intuitive understanding does not exist since the phenomena are not within every-day experience.

1.2 Derivation of Energy Values and Transition Matrix Elements of the Quantum Harmonic Oscillator

1.2.1 From the Classical to the Quantum Oscillator

It is no accident that there are many systems in classical physics that behave as if they were harmonic oscillators. To understand why this is so, consider a particle in the external, one-dimensional potential $U(x)$ depicted in Fig. 1.1. Expanding the potential in a Taylor series about the point of stable equilibrium $x = x_e$,

$$U(x) = U(x_e) + \left.\frac{\mathrm{d}U(x)}{\mathrm{d}x}\right|_{x=x_e}(x - x_e) + \frac{1}{2!}\left.\frac{\mathrm{d}^2U(x)}{\mathrm{d}x^2}\right|_{x=x_e}(x - x_e)^2 + \cdots . \tag{1.2.1}$$

At a point of stable equilibrium x_e, there is a relative minimum of the potential so $\frac{\mathrm{d}U(x)}{\mathrm{d}x}$ is zero at $x = x_e$. Therefore, except for the constant term $U(x_e)$ that shifts all energy levels by the same amount $U(x_e)$, from (1.2.1) it follows that for sufficiently small oscillations the potential energy is proportional to the square of the displacement from equilibrium. Therefore *any* system near equilibrium acts as if it were a harmonic oscillator.

Classically, a harmonic oscillator consists of two objects bound by an attractive force proportional to the relative displacement of the objects. An example is a physical system consisting of masses $m^{(1)}$ and $m^{(2)}$ connected by a spring with a

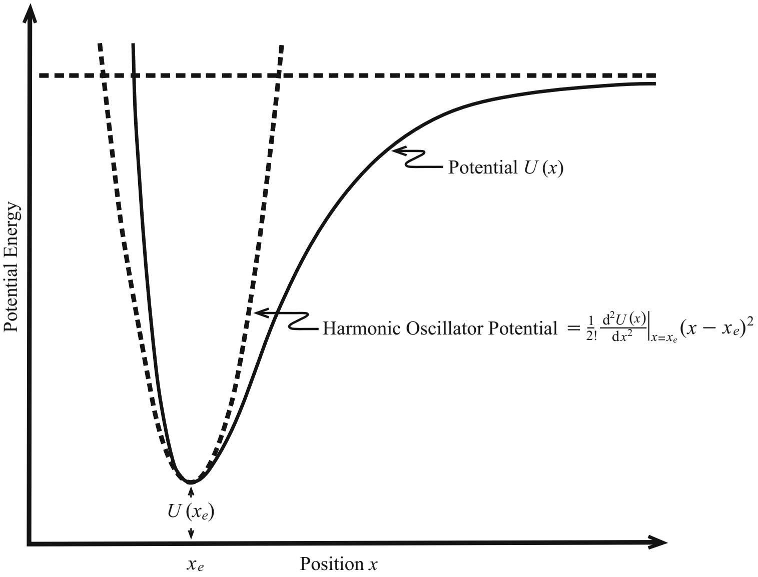


Fig. 1.1 Potential $U(x)$ and the corresponding approximate harmonic oscillator potential

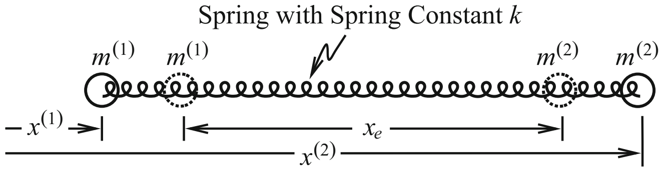


Fig. 1.2 The diatomic molecule as a harmonic oscillator

spring constant k and sliding on a level, frictionless surface. The classical harmonic oscillator is an idealized model (the elastic limit of the spring is not exceeded, there is no friction, etc.) with the characteristic property that it vibrates harmonically.

A quantum harmonic oscillator is a micro-physical system in nature with a mathematical structure that is obtained from the mathematical structure of the classical harmonic oscillator via the Bohr correspondence principle. There are many such micro-physical systems. One example is the diatomic molecule, provided the energy is not too great and there is no rotation. Rather than analyze a system such as a diatomic molecule in terms of its many constituent particles, it is often much easier to analyze the system in terms of how it moves as a whole—such as vibrate or rotate. Motion of the system as a whole is called *collective motion*.

For small oscillations a non-rotating diatomic molecule can be described as if it were a harmonic oscillator, which is the much simpler system depicted in Fig. 1.2.

Denote the positions of the nuclei with masses $m^{(1)}$ and $m^{(2)}$ by $x^{(1)}$ and $x^{(2)}$, respectively, and let the spring have an equilibrium length x_e . Assuming, as shown in Fig. 1.2, that $x^{(2)} > x^{(1)}$, the internuclear separation is given by $x^{(2)} - x^{(1)}$, and the distance the spring is stretched is $x^{(2)} - x^{(1)} - x_e$. The potential energy associated with the stretched spring is

$$\text{potential energy} = \frac{1}{2}k(x^{(2)} - x^{(1)} - x_e)^2. \tag{1.2.2}$$

The potential $U(x^{(2)} - x^{(1)})$ of a diatomic molecule and the approximate harmonic potential $k(x^{(2)} - x^{(1)} - x_e)^2/2$ are depicted in Fig. 1.3 on the following page. The total classical energy $E_T^{\text{Classical}}$ is obtained by adding the kinetic energies of $m^{(1)}$ and $m^{(2)}$ to the potential energy,

$$E_T^{\text{Classical}} = \frac{1}{2}m^{(1)}\left(\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t}\right)^2 + \frac{1}{2}m^{(2)}\left(\frac{\mathrm{d}x^{(2)}}{\mathrm{d}t}\right)^2 + \frac{1}{2}k(x^{(2)} - x^{(1)} - x_e)^2. \tag{1.2.3}$$

The expression for $E_T^{\text{Classical}}$ is greatly simplified by rewriting it in terms of the center-of-mass coordinate X defined by

$$X = \frac{1}{m^{(1)} + m^{(2)}}(m^{(1)}x^{(1)} + m^{(2)}x^{(2)}), \tag{1.2.4a}$$

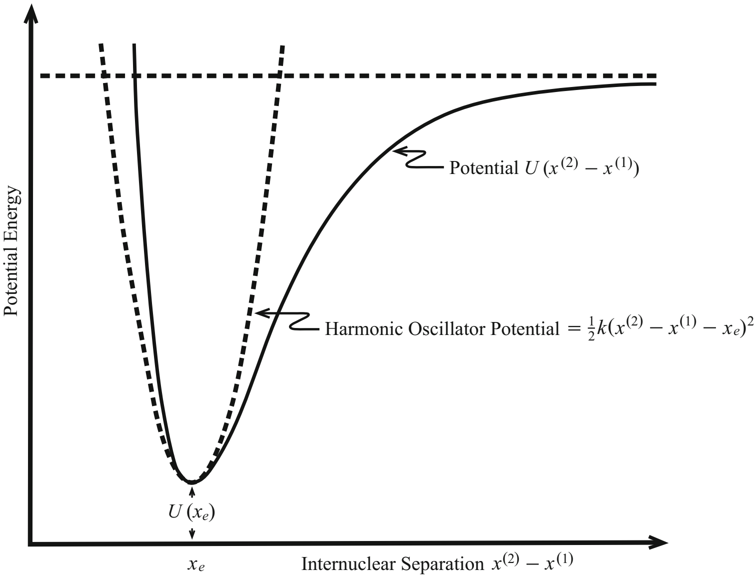


Fig. 1.3 Potential of a diatomic molecule and the corresponding approximate harmonic oscillator potential

and a relative coordinate x defined by

$$x = x^{(2)} - x^{(1)} - x_e. \tag{1.2.4b}$$

In terms of these new variables, the expression for the total classical energy $E_T^{\text{Classical}}$ becomes

$$E_T^{\text{Classical}} = \frac{1}{2}(m^{(1)} + m^{(2)})\left(\frac{\mathrm{d}X}{\mathrm{d}t}\right)^2 + \frac{1}{2}\mu\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \frac{1}{2}kx^2, \tag{1.2.5}$$

where μ is the reduced mass,

$$\mu = \frac{m^{(1)}m^{(2)}}{m^{(1)} + m^{(2)}}. \tag{1.2.6}$$

The first term in (1.2.5) is the kinetic energy associated with the motion of the center of mass. Since there is no potential energy term involving X, the center of mass moves as if it were a free particle with a mass $m^{(1)}+m^{(2)}$. The second and third terms depend on the relative coordinate x. The relative motion of the two particles is the same as that of a single particle with mass μ moving in an external potential

$kx^2/2$. Thus the classical energy $E^{\text{Classical}}$ associated with the relative motion is

$$E^{\text{Classical}} = \frac{1}{2}\mu\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \frac{1}{2}kx^2. \tag{1.2.7}$$

The relative momentum p is given by

$$p = \mu\frac{\mathrm{d}x}{\mathrm{d}t}. \tag{1.2.8}$$

Ignoring the motion of the center of mass, a classical harmonic oscillator is defined as a system with potential energy proportional to the square of the distance from its equilibrium position or as a system with total energy

$$E^{\text{Classical}} = \frac{p^2}{2\mu} + \frac{1}{2}kx^2. \tag{1.2.9}$$

Although the potential depicted in Fig. 1.3 on the preceding page results from the interaction of all the electrons and protons in the diatomic molecule, from the discussion following (1.2.1) it follows that for a non-rotating diatomic molecule undergoing small oscillations, the energy levels can be calculated from the much simpler picture in Fig. 1.2 on page 7, which is that of a diatomic molecule consisting of two nuclei with respective masses $m^{(1)}$ and $m^{(2)}$ bound by an elastic force.

According to the *quantization rules*, the transition from classical to quantum physics is made by replacing classical quantities such as the position x and momentum p by the corresponding quantum mechanical operators Q and P, respectively.[1] The quantum energy operator, the Hamiltonian H, is then obtained from the classical formula for the energy (1.2.9) by making the replacements $E_{\text{classical}} \to H$, $x \to Q$ and $p \to P$. As a consequence, the energy operator for the quantum harmonic oscillator is conjectured to be given by the Hamiltonian

$$H = \frac{P^2}{2\mu} + \frac{1}{2}kQ^2. \tag{1.2.10}$$

---

[1] The basic mathematical notion of quantum mechanics is that of the Hilbert space, where the self-adjoint operators are identified with *observables* and the vectors of the Hilbert space are identified with the *states* of the physical system. In this chapter the mathematical notions of quantum mechanics are introduced gradually in relation to the discussed physical concepts. A systematic, elementary explanation of mathematical ideas of quantum mechanics is discussed in Appendix, page 313. Additional, more complete discussion of precise mathematical formulation of quantum mechanical concepts can be found in the references: A. Bohm, *The Rigged Hilbert Space and Quantum Mechanics*, Lecture Notes in Physics, **78** (1978), Springer-Verlag, Berlin, Heidelberg, New York and A. Bohm and M. Gadella, *Dirac Kets, Gamow Vectors, and Gel'fand Triplets*, Lecture Notes in Physics, **348** (1989), Springer-Verlag, Berlin, Heidelberg, New York.

The operators P and Q are postulated to satisfy the Heisenberg commutation relations

$$[Q, P] \equiv QP - PQ = i\hbar\mathbb{1}\,, \qquad [P, P] = [Q, Q] = 0\,, \tag{1.2.11}$$

where the operator $\mathbb{1}$ is the unit operator and $\hbar = h/2\pi$, where h is Planck's constant (see Eq. (1.1.2)). The constants μ and k are system constants, and their meanings follow from correspondence with the classical system of (1.2.9).

Such a classical system performs oscillations with an angular frequency

$$\omega = \sqrt{\frac{k}{\mu}}\,. \tag{1.2.12}$$

Using (1.2.12) to express k in terms of ω, the quantum system, (1.2.13) can be written as

$$H = \frac{P^2}{2\mu} + \frac{\mu\omega^2}{2}Q^2\,. \tag{1.2.13}$$

In the classical case the energy $E^{\text{Classical}}$, the momentum p, and the position x are real numbers whereas in the quantum-mechanical case the quantities are represented by the self-adjoint[2] operators H, P, and Q, respectively. The constants μ and ω are characteristic of the particular physical system.

1.2.2 Finding the Mathematical Properties of the Operators P, Q, and H

It is easier to perform algebraic manipulations using operators that are dimensionless. To construct the dimensionless operators for the harmonic oscillator, the operators, the available constants, and the dimensions of each are listed in Table 1.1. Note that since $\omega = \sqrt{k/\mu}$, only two of the three constants μ, ω and k are listed. The dimensionless operators that can be constructed from the position operator Q, the momentum operator P, and the Hamiltonian (energy operator) H are, respectively,

$$\sqrt{\frac{\mu\omega}{\hbar}}\,Q\,, \qquad \frac{1}{\sqrt{\mu\omega\hbar}}\,P\,, \qquad \frac{1}{\hbar\omega}\,H\,. \tag{1.2.14}$$

---

[2] To be precise, these operators usually denote essentially self-adjoint operators since they are defined as continuous operators in a dense subspace Φ of the Hilbert space $\mathscr{H}$. Later in this section the difference between $\mathscr{H}$ and Φ will be discussed briefly, but in the remainder of the text, the distinction between $\mathscr{H}$ and its "physical" subspace Φ will not be emphasized.

Table 1.1 The operators and constants for the quantum, harmonic oscillator and the dimensions of each

Operator	Dimension	Constant	Dimension
Q	Length	μ	Mass
P	$\frac{\text{mass} \times \text{length}}{\text{time}}$	ω	$\frac{1}{\text{time}}$
H	$\frac{\text{mass} \times \text{length}^2}{\text{time}^2}$	$\hbar$	$\frac{\text{mass} \times \text{length}^2}{\text{time}}$

It is now a simple matter to express the "dimensionless Hamiltonian" in terms of the "dimensionless position operator" and the "dimensionless momentum operator":

$$\frac{1}{\hbar\omega} H = \frac{1}{\hbar\omega}\left(\frac{P^2}{2\mu} + \frac{k^2}{2} Q^2\right) = \frac{1}{2}\left(\sqrt{\frac{\mu\omega}{\hbar}}\, Q\right)^2 + \frac{1}{2}\left(\frac{P}{\sqrt{\mu\omega\hbar}}\right)^2 . \qquad (1.2.15)$$

In going from the first to the second equality in (1.2.15), the order of the terms has been reversed and the relationship $k = \mu\omega^2$ has been used. To simplify the algebraic manipulations, the "dimensionless Hamiltonian" $H/(\hbar\omega)$ is factored by writing it in terms of new operators a and $a^\dagger$ that are defined in terms of the dimensionless position and momentum operators given in (1.2.15):

$$a \equiv \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\, Q + \frac{i}{\sqrt{\mu\omega\hbar}}\, P\right) , \qquad (1.2.16a)$$

$$a^\dagger \equiv \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\, Q - \frac{i}{\sqrt{\mu\omega\hbar}}\, P\right) . \qquad (1.2.16b)$$

Algebraic manipulations of (1.2.16) yield

$$a^\dagger a = \frac{\mu\omega}{2\hbar} Q^2 + \frac{1}{2\mu\omega\hbar} P^2 - \frac{i}{2\hbar}(PQ - QP) = \frac{1}{\hbar\omega} H - \frac{i}{2\hbar}(PQ - QP) . \qquad (1.2.17)$$

Using the Heisenberg commutation relations (1.2.11), the above equation can be rewritten in the form (Problem 1.1),

$$a^\dagger a = \frac{1}{\hbar\omega} H - \frac{1}{2} 1 \equiv N , \qquad (1.2.18)$$

or

$$H = \hbar\omega(N + \frac{1}{2}) . \qquad (1.2.19)$$

As will soon become apparent, rather than work with the operators H, P, and Q, it is convenient instead to use the operators N, a, and $a^\dagger$. Using the Heisenberg commutation relations (1.2.11), it is straightforward to calculate the commutator of

a and $a^\dagger$ (Problem 1.2)

$$\left[a, a^\dagger\right] = aa^\dagger - a^\dagger a \equiv 1\,. \tag{1.2.20}$$

Because of (1.2.19), calculating all the eigenvalues of N will immediately yield all the eigenvalues of H. It is not obvious that there are any eigenvectors of N (or of H).[3] Here only those representations are considered for which there exists at least one eigenvector $|\lambda\rangle$. This means that an eigenvector $|\lambda\rangle$ is *assumed to exist* that satisfies

$$N|\lambda\rangle = \lambda|\lambda\rangle\,. \tag{1.2.21}$$

As a result of (1.2.19), the eigenvector $|\lambda\rangle$ is also an eigenvector of H with an eigenvalue $\hbar\omega(\lambda + 1/2)$,

$$H|\lambda\rangle = \hbar\omega(N + \frac{1}{2})|\lambda\rangle = \hbar\omega(\lambda + \frac{1}{2})|\lambda\rangle\,. \tag{1.2.22}$$

To indicate that $|\lambda\rangle$ is an eigenvector of both N and the energy H, $|\lambda\rangle$ can be relabeled by the energy E_λ,

$$|\lambda\rangle \equiv |E_\lambda\rangle\,, \qquad E_\lambda = \hbar\omega\left(\lambda + \frac{1}{2}\right)\,. \tag{1.2.23}$$

Equation (1.2.21) only determines $|E_\lambda\rangle$ up to multiplicative constant $\alpha_\lambda \in \mathbb{C}$,[4] where α_λ can depend on λ: the vector $\phi_\lambda \equiv \alpha_\lambda|\lambda\rangle$ is also an eigenvector of N and H with the same eigenvalue because

$$H\phi_\lambda = H\alpha_\lambda|\lambda\rangle = \alpha_\lambda H|\lambda\rangle = \alpha_\lambda E_\lambda|\lambda\rangle = E_\lambda\phi_\lambda\,. \tag{1.2.24}$$

It is customary to work only with normalized eigenvectors so $|\lambda\rangle$ is chosen such that $\langle\lambda|\lambda\rangle = 1$. Then the eigenvectors $|\lambda\rangle$ are determined up to an arbitrary phase factor $e^{i\beta_\lambda}$ where β_λ is real. This phase factor will be chosen by convention.

Starting with the eigenvector $|\lambda\rangle$, additional eigenvectors of N or H are now sought with the help of the relations (Problem 1.3),

$$Na - aN = [N, a] = \left[a^\dagger a, a\right] = a^\dagger\,[a, a] + \left[a^\dagger, a\right]a = -a\,, \tag{1.2.25a}$$

$$\left[N, a^\dagger\right] = \left[a^\dagger a, a^\dagger\right] = a^\dagger\left[a, a^\dagger\right] + \left[a^\dagger, a^\dagger\right] = a^\dagger\,. \tag{1.2.25b}$$

---

[3] In fact there are representations of the commutation relations (1.2.11) in the Hilbert space for which there does not exist even one eigenvector of N.

[4] $\alpha_\lambda \in \mathbb{C}$ means α_λ is a complex number.

The operators a and $a^\dagger$ are first applied to the eigenvector $|\lambda\rangle$ of (1.2.21). Using (1.2.25a),

$$Na|\lambda\rangle = (aN - a)|\lambda\rangle = (a\lambda - a)|\lambda\rangle = (\lambda - 1)\, a|\lambda\rangle\,. \tag{1.2.26a}$$

Similarly, using (1.2.25b) and (1.2.21),

$$Na^\dagger|\lambda\rangle = (a^\dagger N + a^\dagger)|\lambda\rangle = (a^\dagger\lambda + a^\dagger)|\lambda\rangle = (\lambda + 1)\, a^\dagger|\lambda\rangle\,. \tag{1.2.26b}$$

Equation (1.2.26a) reveals that the vector $a|\lambda\rangle$ is an eigenvector of N with an eigenvalue $\lambda - 1$. That is, $a|\lambda\rangle \sim |\lambda - 1\rangle$ unless the vector $a|\lambda\rangle$ is zero. The vector $|\lambda\rangle$ is an eigenvector of N with an eigenvalue λ, and $a|\lambda\rangle$ is an eigenvector of N with an eigenvalue $\lambda - 1$. This means that when the operator a acts on an eigenvector $|\lambda\rangle$, it lowers the eigenvalue of N by one unit.

Similarly, from (1.2.26b), the vector $a^\dagger|\lambda\rangle$ is an eigenvector of N with an eigenvalue $\lambda + 1$, $a^\dagger|\lambda\rangle \sim |\lambda + 1\rangle$, unless the vector $a^\dagger|\lambda\rangle$ is zero. Thus when $a^\dagger$ acts on an eigenvector of N, it increases or raises the eigenvalue of N by one unit. Because a and $a^\dagger$, respectively, decrease and increase the eigenvalues λ of N by 1, $a^\dagger$ is called the raising operator, a is called the lowering operator, and either is called a ladder operator.

The eigenvectors of the operator N are also eigenvectors of H, as can be seen from (1.2.19). The operators a and $a^\dagger$ are applied to the eigenvector $|\lambda\rangle$ an arbitrary number of times. In this way all eigenvalues of N and the Hamiltonian H are found. This task is accomplished by demonstrating the following:

1. The eigenvalues of N are equal to or greater than zero.
2. The smallest eigenvalue of N is zero.
3. The eigenvalues of N increase in integer steps.
4. There is no upper limit to the eigenvalues of N. This means that the space of solutions for the harmonic oscillator Hamiltonian is infinite dimensional.

After establishing the above four statements, it follows that eigenvalues of N are $0, 1, 2, \ldots$.

Step #1 To show that λ satisfies the condition $\lambda \geq 0$, the scalar product of (1.2.21) is taken with $|\lambda\rangle$,

$$\langle\lambda|N|\lambda\rangle = \lambda\langle\lambda|\lambda\rangle = \lambda\, \||\lambda\rangle\|^2\,. \tag{1.2.27a}$$

Since $N = a^\dagger a$,

$$\langle\lambda|N|\lambda\rangle = \langle\lambda|a^\dagger a|\lambda\rangle\,. \tag{1.2.27b}$$

Using the definition of the adjoint operator (see page 321), the above equation becomes

$$\langle\lambda|N|\lambda\rangle = (a|\lambda\rangle, a|\lambda\rangle) = \|a|\lambda\rangle\|^2\,. \tag{1.2.27c}$$

Equating (1.2.27a) and (1.2.27c), the first result is obtained:

$$\lambda = \frac{\|a|\lambda\rangle\|^2}{\||\lambda\rangle\|^2} \geq 0 \;\text{ for all possible } \lambda\,. \tag{1.2.28}$$

The result $\lambda \geq 0$ follows because the norm of any vector is non-negative.

Step #2 The smallest eigenvalue of N is now shown to be $\lambda = 0$ by considering the vector $a|\lambda\rangle$. According to (1.2.26a)

$$Na|\lambda\rangle = (\lambda - 1)\, a|\lambda\rangle\,. \tag{1.2.29}$$

There are two possibilities under which the above equation can be satisfied: Either $a|\lambda\rangle = 0$ or $a|\lambda\rangle$ is an eigenvector of N with an eigenvalue $\lambda - 1$. Suppose that the latter is true. Neglecting for the moment the matter of normalization,

$$|\lambda - 1\rangle \sim a|\lambda\rangle\,. \tag{1.2.30}$$

Repeating the above calculation with $|\lambda - 1\rangle$,

$$|\lambda - 2\rangle \sim a|\lambda - 1\rangle \sim a^2\,|\lambda\rangle\,. \tag{1.2.31}$$

The vector $|\lambda - 2\rangle$ is either zero or an eigenvector of N with eigenvalue $\lambda - 2$. Continuing in this way a sequence of vectors is obtained,

$$|\lambda - n'\rangle \sim (a)^{n'}|E_\lambda\rangle; \quad n' = 0, 1, 2, \ldots \tag{1.2.32}$$

that are eigenvectors of N with eigenvalues $\lambda - n'$ provided none of the eigenvectors $a^{n'}|\lambda\rangle$ are zero. Since, according to (1.2.28), the eigenvalue of N is equal to or greater than zero, this sequence of eigenvectors cannot continue indefinitely. Consequently, there must exist a vector, denoted $|l\rangle$, that is transformed by a into the zero vector 0,

$$a|l\rangle = 0\,, \tag{1.2.33}$$

thus terminating the sequence. Applying the operator $a^\dagger$ to this vector, the vector $|l\rangle$ is seen to be an eigenvector of N with eigenvalue 0 since

$$N|l\rangle = a^\dagger a|l\rangle = a^\dagger 0 = 0 = 0|l\rangle\,. \tag{1.2.34}$$

The vector $|l\rangle = |0\rangle$ is chosen to be normalized. If it initially is not normalized, it will be multiplied by a number C_0 such that the new $|0\rangle$ fulfills the normalization condition

$$\langle 0|0\rangle = 1\,. \tag{1.2.35}$$

Step #3 Starting with the normalized eigenvector $|0\rangle$ of N, $a^\dagger$ is successively applied. Since, according to (1.2.26b), the eigenvalues of N increase in integer steps, the following sequence of vectors,

$$\begin{aligned}
&|0\rangle\,,\\
&|1\rangle = c_1 a^\dagger|0\rangle\,,\\
&|2\rangle = c_2 (a^\dagger)^2|0\rangle\,,\\
&\vdots\\
&|n-1\rangle = c_{n-1}(a^\dagger)^{n-1}|0\rangle\,,\\
&|n\rangle = c_n (a^\dagger)^n|0\rangle = a^\dagger c_n (a^\dagger)^{n-1}|0\rangle = \frac{c_n}{c_{n-1}} a^\dagger|n-1\rangle\,,\\
&\vdots
\end{aligned} \tag{1.2.36}$$

are eigenvectors of N with eigenvalues $0, 1, 2, 3, \ldots$ In (1.2.36) the complex numbers c_n will be chosen below such that $|n\rangle$ is normalized, $\langle n|n\rangle = 1$.

From (1.2.36) and (1.2.26b)

$$N|n\rangle = c_n N a^\dagger (a^\dagger)^{n-1}|0\rangle = \frac{c_n}{c_{n-1}} N a^\dagger|n-1\rangle = n\frac{c_n}{c_{n-1}} a^\dagger|n-1\rangle = n|n\rangle\,. \tag{1.2.37}$$

This means that all $|n\rangle$; $n = 0, 1, 2, \ldots$; are eigenvectors of N with respective eigenvalues n unless one of them $a^\dagger|k\rangle$ is the zero vector, thus terminating the sequence (1.2.36).

Step #4 A proof by contradiction reveals that as n increases, the sequence (1.2.36) never terminates: Suppose that $|k\rangle$ is the last eigenvector so that $a^\dagger|k\rangle = 0$, implying that

$$\|a^\dagger|k\rangle\|^2 = 0\,.$$

Then using (1.2.20), (1.2.18) and (1.2.21)

$$\begin{aligned}
\|a^\dagger|k\rangle\|^2 &= (a^\dagger|k\rangle, a^\dagger|k\rangle) = (|k\rangle, aa^\dagger|k\rangle = (|k\rangle, (a^\dagger a + 1)|k\rangle)\\
&= (|k\rangle, (k+1)|k\rangle) = (k+1)\||k\rangle\|^2 = 0\,.
\end{aligned}$$

But the above result is incorrect because $(k+1)$ and $\||k\rangle\|$ are different from zero. Thus the assumption that the sequence (1.2.36) terminates leads to a contradiction. As a consequence, there is an infinite set of vectors $|n\rangle$, $n = 0, 1, 2, \cdots$. Because each is an eigenvector of the self-adjoint operator H with a different eigenvalue, each eigenvector is orthogonal to all of the others (Problem 1.10),

$$\langle n|m\rangle = 0 \quad \text{for } n \neq m\,. \tag{1.2.38}$$

According to (1.2.22),

$$H|n\rangle = \hbar\omega(N + \frac{1}{2})|n\rangle = \hbar\omega(n + \frac{1}{2})|n\rangle\,. \tag{1.2.39}$$

Therefore, the set of energy eigenvalues of the harmonic oscillator, called the energy spectrum, is

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right), \qquad n = 0, 1, \ldots\,, \tag{1.2.40}$$

as shown in Fig. 1.4. It never terminates.

It is useful and customary to work with eigenvectors of the energy operator (1.2.36) that are normalized. The coefficients $c_n \in \mathbb{C}$ in (1.2.36) are chosen so that

$$\langle n|m\rangle = \delta_{nm} = \begin{cases} 1 & \text{if } \ n = m\,, \\ 0 & \text{if } \ n \neq m\,. \end{cases} \tag{1.2.41}$$

The complex constants c_n are calculated beginning with an equation that follows from (1.2.36),

$$1 = \||n\rangle\|^2 = (c_n\,(a^\dagger)^n|0\rangle, c_n\,(a^\dagger)^n|0\rangle = ((a^\dagger)^n|0\rangle, (a^\dagger)^n|0\rangle|c_n|^2\,. \tag{1.2.42}$$

But from (1.2.36),

$$((a^\dagger)^n|0\rangle, (a^\dagger)^n|0\rangle = \frac{1}{|c_{n-1}|^2}(a^\dagger|n-1\rangle, a^\dagger|n-1\rangle)\,. \tag{1.2.43}$$

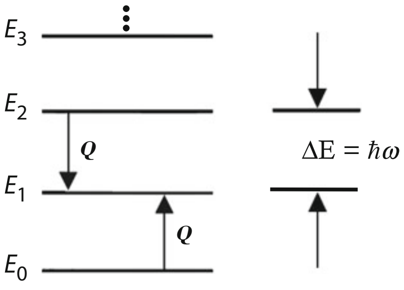


Fig. 1.4 Energy-level diagram of the harmonic oscillator

Substituting the above result into (1.2.42), the following formula is immediately obtained:

$$1 = \frac{|c_n|^2}{|c_{n-1}|^2}(|n-1\rangle, aa^\dagger|n-1\rangle\,. \tag{1.2.44}$$

Using the commutation relations (1.2.20),

$$\begin{aligned}1 &= \frac{|c_n|^2}{|c_{n-1}|^2}(|n-1\rangle, (a^\dagger a+1)|n-1\rangle = \frac{|c_n|^2}{|c_{n-1}|^2}(|n-1\rangle, (N+1)|n-1\rangle \\ &= n\frac{|c_n|^2}{c_{n-1}|^2}(|n-1\rangle, |n-1\rangle = n\frac{|c_n|^2}{|c_{n-1}|^2}\,;\end{aligned} \tag{1.2.45}$$

Hence c_n must be chosen so that

$$n|c_n|^2 = |c_{n-1}|^2\,. \tag{1.2.46}$$

Since $|0\rangle$ is normalized by (1.2.35), $c_0 = 1$, and one solution of (1.2.46) is

$$c_n = \sqrt{\frac{1}{n!}}\,. \tag{1.2.47}$$

There are other solutions of (1.2.46) that differ from (1.2.47) by a phase factor. In fact, the most general solution is $c_n = \frac{1}{\sqrt{n!}}e^{i\alpha_n}$, where α_n is a real constant that can depend on n. It is customary, though not necessary, to choose eigenvectors of H with relative phases that are all zero as the basis system for the space of physical states Φ of the harmonic oscillator. With this choice,

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle\,. \tag{1.2.48}$$

The action of a and $a^\dagger$ on $|n\rangle$ is easily calculated:

$$\begin{aligned}a|n\rangle &= \frac{1}{\sqrt{n!}}a(a^\dagger)^n|0\rangle = \frac{1}{\sqrt{n!}}(a^\dagger a+1)(a^\dagger)^{n-1}|0\rangle \\ &= \frac{1}{\sqrt{n!}}(N+1)\sqrt{(n-1)!}|n-1\rangle = \frac{1}{\sqrt{n!}}(n-1+1)\sqrt{(n-1)!}|n-1\rangle \\ &= \frac{\sqrt{n}}{\sqrt{n}}\frac{\sqrt{n}}{\sqrt{(n-1)!}}\sqrt{(n-1)!}|n-1\rangle = \sqrt{n}|n-1\rangle\,.\end{aligned} \tag{1.2.49}$$

Similarly (Problem 1.11),

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle\,. \tag{1.2.50}$$

From (1.2.16) the operators Q and P can be expressed in terms of a and $a^\dagger$,

$$Q = \sqrt{\frac{\hbar}{2\mu\omega}}(a^\dagger + a)\,, \qquad P = i\sqrt{\frac{\mu\omega\hbar}{2}}(a^\dagger - a)\,. \tag{1.2.51}$$

Example 1.2.1 The operators Q, P and a, $a^\dagger$ (1.2.16) are defined in the abstract vector space, but they can also be defined in the vector space consisting of complex functions $\psi(x)$, which are square integrable $\int |\psi(x)|^2 dx < \infty$. In such a space the action of Q and P on $\psi(x)$ is defined

$$Q\psi(x) = x\psi(x), \quad P\psi(x) = \frac{\hbar}{i}\frac{\partial\psi(x)}{\partial x}.$$

Determine functions $\psi_i(x)$, $i = 0, 1, 2$, which correspond to the energy eigenvectors $|i\rangle$ (1.2.48) of the harmonic oscillator.

Solution The vector $|0\rangle$ fulfills the condition (1.2.33) $a|0\rangle = 0$, so $\psi_0(x)$ fulfills the equation

$$\begin{aligned} 0 = a\psi_0(x) &= \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\, Q\psi_0(x) + \frac{i}{\sqrt{\mu\omega\hbar}}\, P\psi_0(x)\right) \\ &= \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\, x\psi_0(x) + \frac{\hbar}{\sqrt{\mu\omega\hbar}}\frac{\partial\psi_0(x)}{\partial x}\right) \Rightarrow \frac{\mathrm{d}\psi_0(x)}{\mathrm{d}x} = -\frac{\mu\omega}{\hbar}x\psi_0(x). \end{aligned}$$

The solution of the preceding equation for $\psi_0(x)$ is

$$\psi_0(x) = N\mathrm{e}^{-\frac{\mu\omega x^2}{2\hbar}}, \quad N \text{ is the normalization factor.}$$

The normalization factor N is determined from the condition $\int_{-\infty}^{+\infty} |\psi_0(x)|^2\, \mathrm{d}x = 1$ and is equal

$$N = \sqrt[4]{\frac{\mu\omega}{\pi\hbar}} \Rightarrow \psi_0(x) = \sqrt[4]{\frac{\mu\omega}{\pi\hbar}}\mathrm{e}^{-\frac{\mu\omega x^2}{2\hbar}}.$$

From Eq. (1.2.48) we have $|1\rangle = a^\dagger|0\rangle$, so $\psi_1(x)$ is equal

$$\begin{aligned} \psi_1(x) = a^\dagger\psi_0(x) &= \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\, Q\psi_0(x) - \frac{i}{\sqrt{\mu\omega\hbar}}\, P\psi_0(x)\right) \\ &= \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\, x\psi_0(x) - \frac{\hbar}{\sqrt{\mu\omega\hbar}}\frac{\mathrm{d}\psi_0(x)}{\mathrm{d}x}\right) = \sqrt[4]{\frac{4(\mu\omega)^3}{\pi\hbar^3}}x\mathrm{e}^{-\frac{\mu\omega x^2}{2\hbar}} \end{aligned}$$

From the relation $|2\rangle = \frac{1}{\sqrt{2!}}(a^\dagger)^2|0\rangle = \frac{1}{\sqrt{2!}}a^\dagger|1\rangle$ one obtains the function $\psi_2(x)$

$$\psi_2(x) = \frac{1}{2\sqrt{2}}\sqrt[4]{\frac{\mu\omega}{\pi\hbar}}\left(\frac{\mu\omega}{\hbar}x^2 - 1\right)\mathrm{e}^{-\frac{\mu\omega x^2}{2\hbar}}.$$

In the general case the function $\psi_n(x)$ corresponding to the eigenvector $|n\rangle$ is equal

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}}\sqrt[4]{\frac{\mu\omega}{\pi\hbar}}\,H_n\left(\sqrt{\frac{\mu\omega}{\hbar}}\,x\right)\mathrm{e}^{-\frac{\mu\omega x^2}{2\hbar}},$$

where $H_n(z)$ are the n-th order Hermite polynomial of a variable z.

Using (1.2.49)–(1.2.51), it is possible to obtain the matrix element of any power of Q or P between any two energy eigenstates. For example, using the notation $|n\rangle \equiv |E_n\rangle$ one obtains

$$\begin{aligned}\langle E_n|Q|E_m\rangle &= \langle E_n|\sqrt{\frac{\hbar}{2\mu\omega}}(a^\dagger + a)|E_m\rangle \\ &= \sqrt{\frac{\hbar}{2\mu\omega}}\left[\sqrt{m+1}\langle E_n|E_{m+1}\rangle + \sqrt{m}\langle E_n|E_{m-1}\rangle\right] \\ &= \sqrt{\frac{\hbar}{2\mu\omega}}\left[\sqrt{m+1}\,\delta_{n,m+1} + \sqrt{m}\delta_{n,m-1}\right]. \end{aligned} \tag{1.2.52}$$

Two of the non-zero matrix elements of the operator Q are shown by the arrows in Fig. 1.4 on page 16, leading to emission or absorption of dipole radiation.

$$O_{n=2} \to O_{n=1} + \gamma, \qquad \gamma + O_{n=0} \to O_{n=1},$$

where O_n represents an oscillator with quantum number n, and γ denotes a photon with energy $E_\gamma = h\nu = \hbar\omega$.

Placing the matrix elements in a square array,

$$\langle E_n|Q|E_m\rangle = \sqrt{\frac{\hbar}{2\mu\omega}}\begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{1.2.53}$$

Similarly, the matrix of the momentum operator P in the basis $|E_n\rangle$ is

$$\langle E_n|P|E_m\rangle = i\sqrt{\frac{\mu\omega\hbar}{2}}\begin{pmatrix} 0 & -\sqrt{1} & 0 & 0 & \dots \\ \sqrt{1} & 0 & -\sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & -\sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{1.2.54}$$

As mentioned previously, even after the eigenvectors $|E_n\rangle$ are normalized and fulfill (1.2.41), they are determined only up to a factor of modulus 1 that is called a phase factor and usually written as $e^{i\varphi}$ with real "phase angle" φ. Since from $H\,|E_n\rangle = E_n\,|E_n\rangle$ it follows that $H(e^{i\varphi}|E_n\rangle) = E_n(e^{i\varphi}|E_n\rangle)$, there is not a single, normalized eigenvector $|E_n\rangle$ satisfying $H|E_n\rangle = E_n\,|E_n\rangle$ but instead a one-dimensional subspace, denoted $\mathfrak{h}_n$, of eigenvectors with eigenvalue E_n

$$\mathfrak{h}_n = \left\{\phi_n\,\middle|\,\phi_n = e^{i\varphi}\,|E_n\rangle,\ 0 \le \varphi < 2\pi\right\} \quad \text{for every } n = 0, 1, \dots. \tag{1.2.55}$$

These subspaces are orthogonal to each other, written $\mathfrak{h}_n \perp \mathfrak{h}_m$ for $n \neq m$, which means for any $|E_n\rangle \in \mathfrak{h}_n$ and $|E_m\rangle \in \mathfrak{h}_m$ it follows that $\langle E_m|E_n\rangle = 0$.

The set of all linear combinations of $|E_n\rangle$ with components $a_n = \langle E_n|\phi\rangle = (|E_n\rangle, |\phi\rangle)$ that fulfill $\sum |a_n|^2 < \infty$ is the Hilbert space (space of square summable sequences)

$$\mathscr{H} = \left\{\phi|\phi = \sum_{n=0}^{\infty} a_n|E_n\rangle \quad \text{with} \quad a_n \in \mathbb{C} \quad \text{and} \quad \sum_{n=0}^{\infty}|a_n|^2 = (\phi,\phi) < \infty\right\}. \tag{1.2.56}$$

The projection operators[5] $\Lambda_m = |E_m\rangle\langle E_m|$ on the subspaces $\mathfrak{h}_m$ possess the properties

$$\Lambda_m\phi = \sum_{n=0}^{\infty} a_n\,\Lambda_m\,|E_n\rangle = a_m|E_m\rangle\,, \tag{1.2.57}$$

or

$$\Lambda_m\phi_n = \begin{cases}\phi_n & \text{if} \quad m = n\,,\\ 0 & \text{if} \quad m \neq n\,.\end{cases} \tag{1.2.58}$$

The projection operators are also written

$$\Lambda_n = |\phi_n\rangle\langle\phi_n| = e^{i\varphi}|E_n\rangle\langle E_n|e^{-i\varphi} = |E_n\rangle\langle E_n|\,. \tag{1.2.59}$$

---

[5] A selfadjoint operator with the property $\Lambda^2 = \Lambda$ is called a projection operator.

The energy operator H can be expressed in terms of the projection operators Λ_n,

$$H = \sum_0^\infty E_n \Lambda_n = \sum_0^\infty E_n \,|E_n\rangle\langle E_n|\,, \tag{1.2.60}$$

which is called the *spectral representation* of the operator H since the set of eigenvalues E_n is called the spectrum of H.

The subspace $\mathfrak{h}_n$ is, therefore,

$$\mathfrak{h}_n = \{|E_n\rangle\langle E_n|\phi\rangle\,,\ \phi \in \mathscr{H}\}\ , \tag{1.2.61}$$

which is the set of all vectors – normalized or not – that are eigenvectors of H with eigenvalue E_n; $\mathfrak{h}_n$ is called the eigenspace of H, or the energy eigenspace, corresponding to the eigenvalue E_n.

Instead of the space $\mathscr{H}$, the Schwartz space Φ can be considered where

$$\Phi = \left\{\phi|\phi = \sum_{n=0}^\infty a_n|E_n\rangle, a_n \in \mathbb{C},\ \sum_{n=0}^\infty E_n^p|a_n|^2 \equiv \langle\phi|H^p|\phi\rangle < \infty,\ \forall\ p = 0, 1, 2\ldots\right\}. \tag{1.2.62}$$

(the symbol ∀ means “for all”). In $\mathscr{H}$, $(\phi, \phi) < \infty$ and in Φ, $\langle\phi|H^p|\phi\rangle < \infty$. Thus while the space Φ is still an infinite-dimensional space, it is “much smaller” than $\mathscr{H}$,

$$\Phi \subset \mathscr{H}\,, \tag{1.2.63}$$

because the condition on $\phi \in \Phi$ is more restrictive than for $\phi \in \mathscr{H}$. For vectors with a finite number of components, $\phi^{(N)} = \sum_{n=0}^N a_n|E_n\rangle$, this does not make a difference provided all the a_n are finite. But for vectors ϕ with an infinite number of components $a_n = \langle E_n|\phi\rangle \neq 0,\ n = 1, 2, \ldots, \infty$, the spaces in (1.2.56) and (1.2.62) are very different.

The oscillator Hamiltonian (1.2.10), with $k/\mu =$ constant, provides an adequate description in a *limited* energy range. Since vibrating diatomic molecules are only approximately harmonic oscillators described by the Hamiltonian given in (1.2.10), only a finite number of the energy eigenstates $|E_n\rangle$ are of physical importance. Consequently, the question of whether the space Φ of (1.2.62) or the Hilbert spaces $\mathscr{H}$ of (1.2.56) is the correct space of physical states cannot be decided by the oscillator model.

In higher energy ranges, when interactions with other quantum systems and scattering and decay become important, this simple harmonic oscillator model has to be modified. There is some evidence that spaces similar to Φ are better suited to describe observations than Hilbert spaces $\mathscr{H}$ of (1.2.56). The idealization that is chosen here is the Schwartz space Φ because it is the Schwartz space in which all the observables of the harmonic oscillator are represented by *continuous* operators defined in all of Φ.[6] This is not the case for the Hilbert space $\mathscr{H}$ since the operators P and Q that fulfill the fundamental commutation relations (1.2.11) cannot be represented by continuous, Hilbert-space operators defined everywhere in $\mathscr{H}$. This means that in $\mathscr{H}$ the observables P and Q (and many others) cannot form an algebra of observables. That is, the operators cannot be multiplied and added in all of $\mathscr{H}$.

The space of physical states is thus postulated to be a direct sum of one-dimensional energy eigenspaces $\mathfrak{h}_n$ with vectors ϕ that fulfill (1.2.62). It is denoted

$$\Phi = \sum \oplus \mathfrak{h}_n \, . \tag{1.2.64}$$

The quantum harmonic oscillator has a lowest energy value $E_0 = \hbar\omega/2$, known as the zero-point energy. This feature, particular to quantum mechanics, is to be contrasted with the classical convention that the minimum energy of the oscillator is zero. In more complicated molecules that can vibrate with two characteristic frequencies ω and ω', the effect of the zero-point energy can actually be observed when the molecules make a transition from a state with zero-point energy $\hbar\omega'/2$ to a state with a zero-point energy $\hbar\omega/2$.

As shown in Fig. 1.4 on page 16, equidistant energy levels E_n, separated by an amount $\hbar\omega$, are predicted for the *harmonic* oscillator. The energy E_n is associated with a state described by $\Lambda_n = |E_n\rangle\langle E_n|$ or by the one-dimensional subspace $\mathfrak{h}_n$ or by the "ray" $\phi_n = e^{i\varphi}|E_n\rangle$ or by the vector $|E_n\rangle$ up to a phase.

According to (1.2.52), the position operator Q transforms between states of neighboring energy levels,

$$\begin{tikzcd}
 & \mathfrak{h}_{n+1} \\
Q : \mathfrak{h}_n \arrow[ur] \arrow[dr] & \\
 & \mathfrak{h}_{n+1}
\end{tikzcd} \tag{1.2.65}$$

Radiative dipole transitions between neighboring energy levels that involve the absorption of a photon, $O_n + \gamma \to O_{n+1}$, or emission of a photon, $O_n \to O_{n-1} + \gamma$, by the oscillator are caused by the position operator Q. The electric dipole operator of a harmonic oscillator is qQ, where q is a charge characteristic of the particular diatomic molecule. As shall be explained later, the probability of a dipole transition from the state $|E_m\rangle$ to the states $|E_n\rangle$ is proportional to $|\langle E_n|Q|E_m\rangle|^2$ and is, according to (1.2.52), equal to zero except between neighboring energy levels, $n - m = \pm 1$. Thus energy can only be emitted in amounts of $\hbar\omega$, where ω is the angular frequency $\omega = \sqrt{k/\mu}$ of the oscillator with spring constant k and reduced mass μ. Similarly, the energy is also absorbed only in units of $\hbar\omega$. There is only one

---

[6] A. Bohm, *The Rigged Hilbert Space and Quantum Mechanics*, Lecture Notes in Physics, **78** (1978), Springer-Verlag, Berlin, Heidelberg, New York.

spectral line for the harmonically oscillating diatomic molecule with frequency

$$\nu_{nm} = \frac{|E_n - E_{n\pm 1}|}{h} = \frac{|E_n - E_{n\pm 1}|}{2\pi\hbar} = \frac{\omega}{2\pi}|n + \frac{1}{2} - (n \pm 1 + \frac{1}{2})| = \frac{\omega}{2\pi}\,, \tag{1.2.66}$$

where $m = n \pm 1$.

The formula $h\nu = \Delta E$ was Planck's original hypothesis that opened the gate to quantum mechanics. The frequency ν of the blackbody radiation is given by the frequency $\nu = \omega/(2\pi) = \sqrt{k/\mu}/(2\pi)$ of the oscillators in the wall of Planck's black body.

In addition to the frequency $\nu = \omega/(2\pi) = \sqrt{k/\mu}/(2\pi)$ predicted by (1.2.66), which results from transitions between neighboring energy levels, for the HCl molecule radiation at lower intensities is also observed with frequencies that are integer multiples $m = 2, 3, 4$, and 5 of this frequency,

$$\nu_{0m} = m\,\frac{\omega}{2\pi} = m\,\nu\,. \tag{1.2.67}$$

These higher frequencies result from transitions between energy levels that are not adjacent. The transitions for the HCl molecule are shown in Fig. 1.5. The presence of transitions with $n = 2, 3, 4$, and 5 is an indication that the HCl molecule is not precisely a harmonic oscillator defined by (1.2.10). The harmonic oscillator defined by (1.2.10) is just the benchmark for understanding the vibrating molecule: the nature of the vibrating molecule is understood both by its agreement with the harmonic oscillator model and by its deviations from the predictions of the harmonic oscillator model. However, since the intensities of transitions with $m > 1$ rapidly decrease, the harmonic oscillator model already provides a very good understanding in the energy regime $E \approx 0.5\,\text{eV}$.

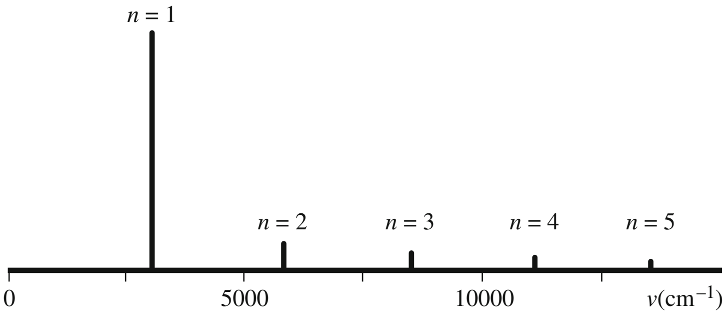


Fig. 1.5 Coarse structure of the infrared spectrum of the diatomic molecule HCl. The intensity actually decreases five times faster than indicated by the height of the vertical lines. Herzberg (1966), vol. 1

In the present section the energy levels of the harmonic oscillator have been determined and the transitions between these energy levels have been discussed. In Sect. 1.4 the scattering of electrons by quantum harmonic oscillators will be used to introduce the fundamental notions of quantum mechanics: states and observables.

1.3 Pure States, Mixtures and Quantum Mechanical Probabilities and Transition Rates

1.3.1 Introduction

From the Hamiltonian (1.2.10), the commutation relations (1.2.11) *and* the assumption that there is at least one energy eigenvector, it was possible to derive the energy spectrum of the harmonic oscillator. The spectrum consists of a set of discrete, equally-spaced energy eigenvalues as shown in Fig. 1.4 on page 16. The spacing between adjacent energy levels is $E_{n+1} - E_n = \hbar\omega$, where $\omega = \sqrt{k/\mu}$ is the angular frequency with which the molecule vibrates, and the system constants k and μ are the spring constant and reduced mass, respectively.

Physical quantum harmonic oscillators occur in nature in many different forms, in particular as diatomic molecules such as O_2, HCl, and CO. Each of these oscillators has different values of the system constants k and μ, resulting in a different value of angular frequency ω. But otherwise the Hamiltonian of each of the vibrating diatomic molecules has the same form given by (1.2.10) leading to equal spacing between energy levels. The commutation relations (1.2.11) are independent of the particular physical system since they contain only the universal constant $\hbar$. Therefore, one expects to find many physical systems for which the energy levels are equally spaced or approximately equally spaced as predicted by (1.2.40).

As is generally the case, models such as the oscillator or the harmonically vibrating dumbbell are applicable only in a limited energy range. In lower energy ranges $\Delta E < \hbar\omega$, the CO molecule does not behave like a vibrator, but it can still perform rotations so it behaves as a rotator. In the energy range 0.1 eV-1.0 eV the harmonic oscillator provides a good model of the CO molecule: there are only small deviations from a harmonic force, and the lower, adjacent energy levels are equally spaced. In higher energy ranges the electronic structure of the atom becomes visible, and the energy spectra have spacings more similar to atomic spectra.

Two different processes for observing the energy levels of the vibrating CO molecule will now be discussed: energy-loss experiments and radiative transitions. Each of the two processes provides an opportunity for measuring both the spacing between energy levels and the probability of transition between two energy levels.

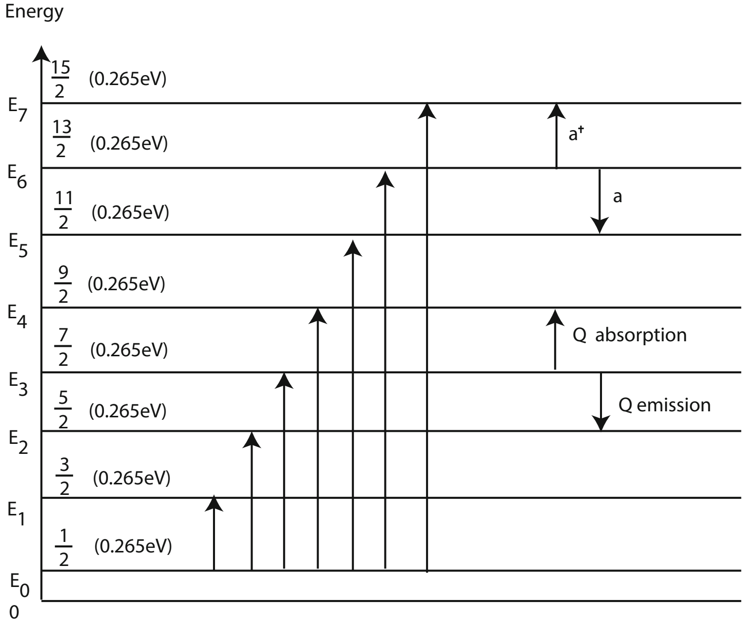


Fig. 1.6 Energy levels of the CO molecule. Energy transitions are indicated by arrows

Transitions in energy-loss experiments result from the collision of an electron beam with the oscillators; Intensities are not governed by the dipole matrix element $\langle E_{n'}|Q|E_n\rangle$. As a result, scattering electrons off of CO molecules can lead to transitions between various energy levels. This is shown in Fig. 1.6 by the arrows from the ground state E_0 to the seven excited levels $E_1, E_2, \ldots, E_7$. The discussion of this experiment will also be used to introduce the fundamental concepts of quantum mechanics: the quantum mechanical state, the observable, and the quantum mechanical probabilities.

The radiative transitions, which will be discussed in Sect. 1.5, are predominantly dipole transitions with the intensity proportional to the square of the magnitude of matrix element of the position operator $|\langle E_{n'}|Q|E_n\rangle|^2$; therefore, from (1.2.52) transitions can occur only between neighboring energy levels as shown by the arrows for absorption and emission in Fig. 1.6. The radiative dipole transitions due to the absorption and emission of electromagnetic radiation of photons occur with the energy $E_\nu = 0.265\,\text{eV}$ or with frequency

$$\nu = \frac{|E_n - E_{n\pm 1}|}{2\pi\hbar} = \frac{0.265\,\text{eV}}{2\pi(6.58\times 10^{-16}\,\text{eV}\cdot\text{s})} = 6.41\times 10^{13}\,\text{s}^{-1}\,, \qquad (1.3.1)$$

and are in the infrared region.

1.3.2 Energy-Loss (Franck-Hertz) Experiments

In energy-loss experiments[7] the energy is measured that is lost by an electron e in a collision with an oscillator O_0 in its ground state,

$$e + O_0 \longrightarrow e' + O_n \quad n = 0, 1, 2, \ldots, 7\,. \tag{1.3.2}$$

The electron with initial energy E_e collides with an oscillator in its ground state that has energy E_0. The oscillator is excited into a vibrational state O_n with energy $E_n,\ n = 0, 1, 2, \ldots$, and the electron has final energy $E_{e'}$. Since energy is conserved in the collision,

$$E_e + E_0 = E_{e'} + E_n\,. \tag{1.3.3}$$

The energy $E_e - E_{e'}$ lost by the electron during the collision is, according to (1.2.40), predicted to be a multiple n of $\hbar\omega$:

$$E_e - E_{e'} = E_n - E_0 = \varDelta E_n = \hbar\omega\left(n + \frac{1}{2}\right) - \hbar\omega\left(\frac{1}{2}\right) = n\hbar\omega\,, \tag{1.3.4}$$

$$\omega = \sqrt{\frac{k}{\mu}}, \quad n = 0, 1, 2, \ldots.$$

It is possible to determine the spacing $E_n - E_0$ between energy levels of the oscillator and verify that the spacing between adjacent energy levels is equidistant as predicted by (1.2.40). Figure 1.6 on the previous page shows the energy levels (1.2.40) predicted for the vibrating CO molecule. Adjacent energy levels are equidistant, and the spacing $E_{n+1} - E_n = \hbar\omega$ can be determined by measuring $E_e - E_{e'}$ in an energy-loss experiment.

The schematic diagram for such an experiment is given in Fig. 1.7a. A beam of electrons leaves a monochromator with energy in a very narrow energy range centered around E_e. The electrons enter a collision chamber with a molecular beam of CO molecules in the ground state $O_{n=0}$, which is an ensemble of CO molecules kept at a low temperature so that the molecules are in their vibrational ground state and moving in the vertical direction of Fig. 1.7b. Some of the electrons scatter into an analyzer that focuses only electrons with an energy $E_{e'}$ onto the detector. In the specific experiment described in Fig. 1.7 on the facing page, the energy resolution is 0.06 eV. The energy $E_{e'}$ selected by the analyzer can be varied, allowing the measurement of the intensity I (the electron current at the detector) of the electrons as a function of the energy $\varDelta E = E_e - E_{e'}$ lost by the electron. According to (1.3.4) the energy lost by the electron equals the energy $\varDelta E_n$ transferred to the CO molecule.

---

[7] From G. J. Schultz, *Phys. Rev.* **135**, A998 (1964), with permission.

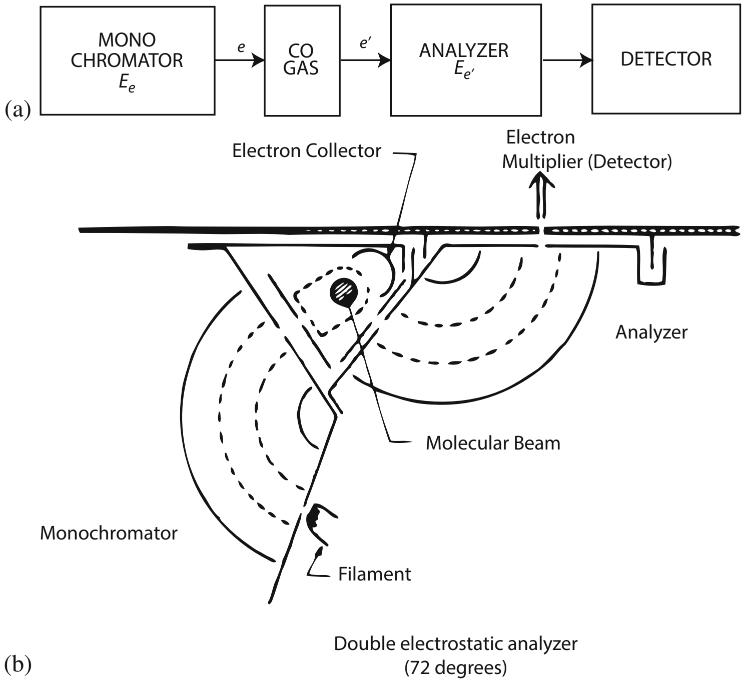


Fig. 1.7 (**a**) Schematic diagram of an energy-loss experiment. (**b**) Schematic diagram of a double electrostatic analyzer. Electrons are emitted from the thoria-coated iridium filament. They then pass between the cylindrical grids at an energy of about 2.05 eV and are accelerated into a collision chamber where they are crossed with a molecular beam. Those electrons scattered into the acceptance angle of the second electrostatic analyzer pass between the cylindrical grids at an energy from 0 to approximately 2 eV. The electrons pass the exit slit into the second chamber and impinge on an electron multiplier [from G.J. Schulz, *Vibrational Excitation of* N_2, *CO, and* H_2 *by Electron Impact*, Phys. Rev. **135**, A988, 1964, with permission]

The results observed in an actual experiment,[8] performed with the apparatus depicted schematically in Fig. 1.7, is shown in Fig. 1.8 on the following page. A maximum intensity in the detected electron current in Fig. 1.7 occurs for an energy loss $\Delta E_0 = E_e - E_{e'} = 0$, implying that a major fraction of the electrons in the current are scattered elastically and do not lose any energy. This is shown by the first maximum in Fig. 1.8 on the following page where the relative intensity has been rescaled by a factor of $\frac{1}{3}$. A second relative maximum of intensity occurs for electrons that have lost energy $\Delta E_1 \simeq 0.265\,\text{eV}$, demonstrating that a portion of the electrons with energy E_e do in fact transfer energy in the amount ΔE_1 to the CO molecules. The third peak occurs for an energy loss $\Delta E_2 = 2E_1$, and so forth.

---

[8] The first experiment of this kind was performed by James Franck and Gustav Hertz; Verhandlung Dtsch. Physikalischen Gesellschaft **16** 457 (1914). The experiment discussed here is by Schulz, G. J.; Phys. Rev. **B5** A988 (1964).

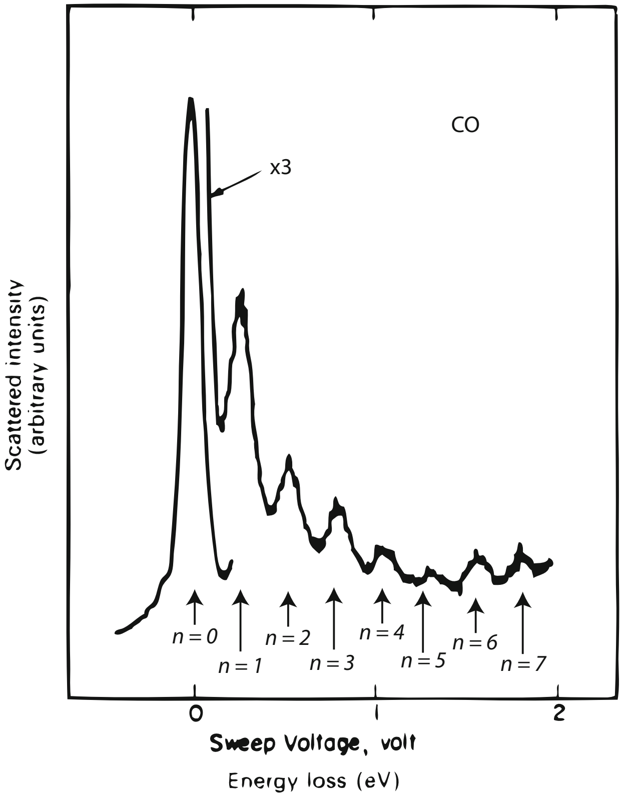


Fig. 1.8 Energy spectrum of electrons with an incident energy of 2.05 eV that are scattered from CO [from G.J. Schulz, *Vibrational Excitation of* N_2*, CO, and* H_2 *by Electron Impact*, Phys. Rev. **135**, A988, 1964, with permission]

Figure 1.8 reveals that the scattered electron current is comprised of electrons that have transferred one of the eight discrete amounts of energy $\Delta E_0, \Delta E_1, \ldots, \Delta E_7$ to the CO molecules. After the collision these CO molecules are in the ground state with energy E_0 and in excited energy levels with energies E_1, E_2,..., E_7.

From the experimental data[9] it follows that the energy lost by electrons is discrete. Thus it is possible to conclude that CO molecules cannot be excited to any arbitrary energy. Only a discrete number of energy values E_n are possible, revealing experimentally that the diatomic molecule CO has discrete, equally-spaced energy levels E_n, as predicted by (1.2.40). The experimental data indicate that the CO molecule is a harmonic oscillator in this energy range. The experimentally-determined energy spectrum is represented by the energy level diagram in Fig. 1.6 on page 25.

---

[9] Multiple scattering of electrons and CO molecules is negligible because the intensity of the electron current and the density of the CO molecules are sufficiently low.

1.3.3 The State of the Ensemble of CO Molecules Participating in the Energy-Loss Experiment

In the scattering process (1.3.2), the interaction between the electrons and the molecule $O_{n=0}$ of the molecular beam leads to the following transitions:

$$O_o \to O_n \begin{cases} n = 0 & \text{for elastic scattering} \\ n = 1, 2, \ldots, 7 & \text{for transitions to excited states as indicated} \\ & \text{by arrows in Fig. 1.6 on page 25.} \end{cases}$$

When the final energy of the oscillator is the same as its initial energy E_0, the collision is said to be elastic. When the final energy of the oscillator is $E_n,\ n \geq 1$, kinetic energy of an electron has been transferred to the oscillator, and the collision is called inelastic.

The intensity of the transition from O_0 to the various excited levels O_n can be measured as function of the energy lost by the electron,

$$\text{Energy loss } = E_e - E_{e'}\,,$$

where E_e is the energy of the monochromator setting and $E_{e'}$ is the energy of the analyzer setting in the experiment of Fig. 1.7 on page 27. As shown in Fig. 1.8 on the facing page, the intensity of the transitions measured by the detector—and therefore the number of CO molecules that participated in these transitions—has a set of 7 discrete peaks for which the energy loss takes the set of discrete values

$$E_e - E_{e'} = E_n - E_0, \quad n = 0,\ n = 1, 2, \ldots, 7.$$

For $n = 0$, the process (1.3.2) is an "elastic collision", because

$$E_e + E_0 = E_{e'} + E_0\,,$$

and there is no change in the CO molecule's energy. For $n > 0$, the process (1.3.2) is an "inelastic collision,"

$$E_e + E_0 = E_{e'} + E_n, \quad n \neq 0\,,$$

because part of the kinetic energy $E_e - E_{e'}$ has been "lost" to the intrinsic energy $E_n - E_0$ of the CO molecule.[10] After a collision an electron has lost energy, and the

---

[10] The kinetic energy of the CO molecules moving in the perpendicular direction in the molecular beam of Fig. 1.7 on page 27 is of a different order of magnitude because $m_{\text{CO}} \gg m_e$.

Table 1.2 For the peaks in Fig. 1.8 on page 28, the energy associated with each peak (or state), the height of or number of molecules in each peak, the state associated with each peak, and the energy associated with each peak

Peaks at $E_e - E_{e'}$	$E_0 - E_0$	$E_1 - E_0$	$E_2 - E_0$	…	$E_n - E_0$
N_n molecules have been excited into the n-th vibrational state	$h_0 \sim N_0$ $\vert 0\rangle\langle 0\vert$	$h_1 \sim N_1$ $\vert 1\rangle\langle 1\vert$	$h_2 \sim N_2$ $\vert 2\rangle\langle 2\vert$		$h_n \sim N_n$ $\vert n\rangle\langle n\vert = \vert E_n\rangle\langle E_n\vert$
of energy	E_0	E_1	E_2		E_n

energy loss is used to excite a CO molecule into a higher vibrational state O_n,

$$\text{energy loss } \equiv E_e - E_{e'} = E_n - E_0.$$

The experiment[7] in Fig. 1.7 on page 27 produced peaks in the electron current registered in the detector at seven different values of energy loss. The height h_n of the n-th peak is shown in Fig. 1.8 on page 28. The interpretation is the following: At the i-th peak with energy loss $E_i - E_0$, a number N_i of molecules have been excited into the i-th energy level. This is shown in detail in Table 1.2.

From the height h_n of the n-th peak in Fig. 1.8 on page 28, it follows that the fraction w_n of molecules in each peak is

$$w_n = \frac{h_n}{h}, \tag{1.3.5}$$

where h is he sum of the heights of the eight peaks,

$$h = \sum h_n. \tag{1.3.6}$$

Letting N_n be the number of CO molecules that have energy E_n, then w_n is also given by

$$w_n = \frac{N_n}{N}, \tag{1.3.7}$$

where N is the total number of molecules in all eight peaks,

$$N = \sum N_n. \tag{1.3.8}$$

The number N of CO molecules that participate in this energy-loss experiment is typically a large number on the order of Avogadro's number $N_A = 6.022 \times 10^{23}$. But in some experiments the number of members in a quantum mechanical ensemble may be orders of magnitude smaller. Neither N nor N_n can be counted since the electron detector measures the current, but h_n and N_n are proportional to the current.

From (1.3.6) and (1.3.8), the w_n fulfill

$$\sum w_n = \sum \frac{N_n}{N} = \sum \frac{h_n}{h} = 1\,. \tag{1.3.9}$$

After the collisions each of the N_n molecules has energy $E_n = E_0 + n \times (0.265)$ eV. The molecules are said to be in the n-th energy level associated with an eigenvector $|E_n\rangle$ of the Hamiltonian H. The height h_n of each peak is proportional to the number of electrons N_n detected in the n-th peak, so the height h_n is proportional to the number of CO molecules in the "excited state" $|E_n\rangle\langle E_n|$. The numbers w_n are the ratios of two large numbers, $w_n = N_n/N$, that are measured as intensities of the electron current at the electron detector in Fig. 1.7b on page 27. It is reasonable to think that the w_n are proportional to the number of electrons that have lost energy $\Delta E_n = E_n - E_0$, $n = 0, 1, \ldots, 7$ to the molecules. Specifically, w_n, which is proportional to the intensity of the molecules scattered with an energy E_n as shown in Fig. 1.8 on page 28, is the probability that the CO molecule is in the n-th energy level,

$$w_n = \frac{h_n}{h} = \frac{N_n}{N}\,. \tag{1.3.10}$$

Equivalently, w_n is the *probability* of finding a CO molecule in a scattered molecular beam that has acquired the energy $E_n - E_0$ from the scattered electrons.

1.3.4 Interpretation of the Experimental Results

In the energy-loss experiment the energy E_0 is obtained N_0 times, the energy E_1 is obtained N_1 times, ..., and the energy E_7 is obtained N_7 times. There are a total of N collision events where $N = N_0 + N_1 + \cdots + N_7$. For the N measurements, the average value for energy is denoted $\langle E\rangle$ and is given by

$$\langle E\rangle = \frac{1}{N}(N_0E_0 + N_1E_1 + \cdots + N_7E_7) = \sum_{n=0}^{7} \frac{N_n}{N} E_n = \sum_{n=0}^{7} w_n\, E_n\,. \tag{1.3.11}$$

The operators $|E_n\rangle\langle E_n|$ or the eigenvectors $|E_n\rangle$ represent the n-th energy eigenstate, and w_n is the probability that the CO molecule has been excited into the n-th vibrational level. With the numbers w_n and the projection operators $\Lambda_n = |E_n\rangle\langle E_n|$ of (1.2.57)–(1.2.59), an operator ρ is defined by

$$\rho = \sum_{n=0}^{7} w_n |E_n\rangle\langle E_n|\,. \tag{1.3.12}$$

This operator ρ represents the **state** of all CO molecules that participated in the energy-loss experiment of Fig. 1.7 on page 27. The operator ρ, called statistical operator or density operator, describes the state of the ensemble of molecules (all the CO molecules that participate in this scattering experiment).

Example 1.3.1 Calculate the trace of the statistical operator ρ.[11]

Solution

$$\operatorname{Tr}\rho \equiv \sum_i \langle E_i|\rho|E_i\rangle = \sum_i \langle E_i| \sum_{n=0}^{7} w_n |E_n\rangle\langle E_n|E_i\rangle \,.$$

Since the vectors $|E_i\rangle$ are an orthonormal basis satisfying $\langle E_n|E_i\rangle = \delta_{ni}$,

$$\operatorname{Tr}\rho = \sum_i \sum_n w_n \delta_{in}\delta_{in} = \sum_i w_i = 1 \,.$$

The last equality follows from (1.3.9).

The number N_n is the number of molecules that have been excited into the n-th energy level; N_n/N is the fraction of molecules that have been excited into the n-th level of energy E_n. The ratio $w_n = N_n/N$ is thus the probability of finding a molecule that has been excited into the n-th energy level from among of all (participating) molecules in the ensemble.

As will now be shown, the probability w_n for finding the value E_n in the state ρ of (1.3.12) can be calculated as the trace of the operator $(\Lambda_n\rho)$, which is called "expectation value of the observable Λ_n in the state ρ." Calculating this trace using the complete system of energy eigenvectors $|E_i\rangle$,

$$\operatorname{Tr}(\Lambda_n\rho) = \operatorname{Tr}(|E_n\rangle\langle E_n|\rho) = \sum_{i=1}^{\infty} \underbrace{\langle E_i|E_n\rangle}_{\delta_{in}} \langle E_n|\rho|E_i\rangle = \langle E_n|\rho|E_n\rangle = w_n \,, \tag{1.3.13}$$

where (1.3.12) has been used to obtain the last equality.

---

[11] The trace $\operatorname{Tr} A$ of an $N \times N$ matrix A_{ij} is defined by

$$\operatorname{Tr} A = \sum_{i=1}^{N} A_{ii} \,.$$

Similarly, the trace $\operatorname{Tr} O$ of an operator O is defined to be sum over the diagonal matrix elements of the operator in any basis:

$$\operatorname{Tr} O = \sum_{\nu=0}^{\infty} (\phi_\nu, O\phi_\nu) \,,$$

where ϕ_ν is any basis system of the (infinite dimensional) space Φ. Operators O, which fulfill the condition $\sum_{\nu=0}^{\infty}(\phi_\nu, (O^\dagger O)^{\frac{1}{2}}\phi_\nu) < \infty$, are called trace-class operators. Here $(O^\dagger O)^{\frac{1}{2}}$ is the non negative, unique square root of $O^\dagger O$.

Thus,

$$(\text{Expectation value of observable } \Lambda_n \text{ in state } \rho) \equiv \mathrm{Tr}(\rho\Lambda_n) = w_n\,. \qquad (1.3.14)$$

The "expectation value of the observable H in the state ρ" is denoted by $\langle H\rangle_\rho$ and is defined by

$$\langle H\rangle_\rho \equiv \mathrm{Tr}(H\rho)\,. \qquad (1.3.15a)$$

The calculation of the expectation value of H in the state (1.3.12) gives

$$\begin{aligned}\langle H\rangle_\rho \equiv \mathrm{Tr}(H\rho) &= \sum_i \langle E_i|\rho H|E_i\rangle = \sum_i w_i \langle E_i|H|E_i\rangle \\ &= \sum_i w_i E_i = \sum_i \frac{N_i}{N} E_i \equiv \langle E\rangle\,. \end{aligned} \qquad (1.3.15b)$$

Thus the calculated quantity "the expectation value of the observable H in the state ρ"$\equiv$ $\mathrm{Tr}(H\rho)$ is identified with the weighted average of the energy values $E_i = 0.265 \cdot (i + \frac{1}{2})$ eV, $i = 0, 1, 2 \ldots$ measured by the energy-loss experiment and calculated in (1.3.11) :

$$\left\{\begin{matrix}\text{expectation value calcu-}\\ \text{lated in the theory}\end{matrix}\right\} \doteqdot \left\{\begin{matrix}\text{observed average value}\\ \text{of the experiment}\end{matrix}\right\}. \qquad (1.3.16a)$$

Expressed as symbols, the above equation becomes

$$\langle H\rangle_\rho \equiv \mathrm{Tr}(H\rho) \doteqdot \langle E\rangle = \sum_i \frac{N_i}{N} E_i = \sum_i w_i E_i\,. \qquad (1.3.16b)$$

The symbol $\doteqdot$ in the above equation expresses the correspondence between the prediction $\mathrm{Tr}(H\rho)$ of the theory and the experimental average $\langle E\rangle$ and indicates that these quantities should be equal to within experimental errors.

The calculated expectation value $\mathrm{Tr}(|E_n\rangle\langle E_n|\rho)$ of the observable $\Lambda_n \equiv |E_n\rangle\langle E_n|$ in the state ρ is – according to (1.3.13) – the observed probability $w_n = N_n/N$ given by the ratio of large numbers N_n/N, where N_n can be thought of as the number of CO molecules in the n-th energy level, and N is the total number of CO molecules that take part in the energy-loss experiment.

What is the expectation value of an observable such as P, Q, or P^2 in the energy eigenstate described by a vectors $\phi_{E_n} = |E_n\rangle e^{i\varphi}$ or by the operator $|\phi_n\rangle\langle\phi_n| = |E_n\rangle\langle E_n| = \Lambda_n$? In general the expectation value of an observable A in the energy eigenstate $\Lambda_n = |\phi_n\rangle\langle\phi_n|$, often denoted as $\langle A\rangle_{\phi_n}$, is

$$\langle A\rangle_{\phi_n} \equiv \mathrm{Tr}(|\phi_n\rangle\langle\phi_n|A) = \sum_i \langle E_i|\Lambda_n A|E_i\rangle = \langle E_n|A|E_n\rangle\,. \qquad (1.3.17)$$

Specifically, the expectation value of the energy operator H in the energy eigenstate $|\phi_n\rangle\langle\phi_n| = \Lambda_n$ is the energy eigenvalue

$$\langle H\rangle_{\phi_n} \equiv \mathrm{Tr}(|E_n\rangle\langle E_n|H) = E_n\,\mathrm{Tr}(|E_n\rangle\langle E_n|) = E_n\,. \tag{1.3.18}$$

This means that the measurement of the observable H in one of its eigenstates ϕ_n gives the eigenvalue E_n of H in this state.

As a generalization of the expectation value in (1.3.17) to the expectation value of an arbitrary observable A such as P, Q, P^2, or $f(P, Q)$ in the general state ρ, the following **Axiom** is conjectured:

$$\left\{\begin{matrix}\text{Expectation value of observ-}\\ \text{able } A \text{ in a state } \rho\end{matrix}\right\} \equiv \langle A\rangle_\rho = \mathrm{Tr}(A\rho) = \mathrm{Tr}(\rho A)\,. \tag{1.3.19}$$

Applying this axiom to the mixed state ρ in (1.3.12),

$$\begin{aligned}\mathrm{Tr}(A\rho) &= \sum_{n=0}^{7} w_n\,\mathrm{Tr}(A|E_n\rangle\langle E_n|) = \sum_m\sum_{n=0}^{7} w_n\langle E_m|A|E_n\rangle\langle E_n|E_m\rangle\\ &= \sum_m\sum_{n=0}^{7} w_n\langle E_m|A|E_n\rangle\delta_{n,m} = \sum_n w_n\langle E_n|A|E_n\rangle\,. \end{aligned}\tag{1.3.20}$$

The above equation shows that the expectation value of an observable A is the weighted average of the expectation values $\langle E_n|A|E_n\rangle$ of A in the energy eigenstates $|E_n\rangle\langle E_n|$.

1.4 Fundamental Postulates of Quantum Mechanics

For a theory the fundamental postulates such as (1.3.19) cannot be derived. Using insight provided by many experimental results, the fundamental postulates can only be conjectured. The energy-loss experiment discussed in Sect. 1.3.2 is used to justify the following fundamental postulates of quantum mechanics, called axioms:

1.4.1 Algebra of Observables—Fundamental Postulate I

A physical observable, defined by the process by which it is measured, is represented by a linear operator A in a linear space. The mathematical image of a physical system is an algebra of linear operators in a linear, scalar-product space Φ.[12] The algebra is generated by fundamental observables representing basic physical quantities, and multiplication is defined by algebraic relations among these fundamental observables.

For the simple example of the one-dimensional harmonic oscillator, the algebra of observables is generated by the following (essentially self-adjoint) operators in the space Φ:

H represents the observable energy
P represents the observable momentum
Q represents the observable position

The defining algebraic relation that is fulfilled by these operators is the canonical commutation relation that is universally satisfied in non-relativistic quantum mechanics,

$$[Q, P] \equiv QP - PQ = i\hbar\mathbb{1}\,, \quad [P, P] = [Q, Q] = 0\,, \tag{1.4.1}$$

where $\hbar$ is a universal constant and $\mathbb{1}$ is the unit operator. The Hamiltonian for the harmonic oscillator is

$$H = \frac{P^2}{2\mu} + \frac{\mu\omega^2}{2}Q^2\,, \tag{1.4.2}$$

where μ and ω are system constants.

1.4.2 Quantum Mechanical States—Fundamental Postulate II

The state of the quantum system is represented by a positive, hermitian, trace-class operator ρ, which is called the *statistical or density operator*. The expectation value of an observable A in the state ρ is given by

$$\langle A\rangle_\rho = \mathrm{Tr}(A\rho) \quad \text{where} \quad \mathrm{Tr}\,\rho = 1\,. \tag{1.4.3}$$

The "expected value" or "expectation value" of the observable A in the state ρ, denoted $\langle A\rangle_\rho$ or just $\langle A\rangle$, is determined by experimental procedures that prepare the state ρ as well as other experimental procedures that define the observable A.

---

[12] Here, for the description of the spectrum and structure of quantum systems such as the oscillator, the space Φ can be a Schwartz space as given in (1.2.62). When discussing scattering, resonance, and decay phenomena, Φ must be a pair of Hardy spaces, one for states and the other for observables.

A series of N measurements of the observable A are carried out to determine its "expectation value" $\langle A\rangle$. Either the measurement is performed on N identical quantum systems or the measurement is performed N times on one physical system while making certain that the system is identically prepared for each measurement.

If the numbers measured for the observable A are denoted $a_0, a_1, \ldots, a_n, \ldots,$ and N_n is the number of measurements that gave the result a_n, then $\langle A\rangle$ is given by the average value of these measurements,

$$\langle A\rangle = \sum_n \frac{N_n}{N} a_n\,, \tag{1.4.4}$$

for sufficiently large N. That is, N must be large enough that the right-hand side of (1.4.4) would fluctuate only slightly if N were increased. The ratio $w_n = N_n/N$ is the probability for obtaining the value a_n in one measurement. Thus $\langle A\rangle$ is the experimentally determined average value.

The operator A is the mathematical representation of an observable, and ρ is the mathematical representation for a state. If the mathematical image of the quantum system (i.e. the theory for the system) is known, then it is possible to calculate $\mathrm{Tr}(A\rho)$, the mathematical quantity that follows from the theory. The average value $\langle A\rangle$ in (1.4.4) is the number measured in experiments. The equality

$$\langle A\rangle \equiv \sum_n \frac{N_n}{N} a_n \doteqdot \mathrm{Tr}(A\rho) = \langle A\rangle_\rho\,, \tag{1.4.5}$$

expresses the equality between the value of $\langle A\rangle$ measured by a large number N of experimental data and the quantity $\langle A\rangle_\rho$ calculated by the mathematical theory. As mentioned in the previous section, the symbol $\doteqdot$ in (1.4.5) indicates that the experimental values $\langle A\rangle$ and the theoretical predictions $\mathrm{Tr}(A\rho)$ should be equal to within experimental errors.

The state operator (also called the statistical operator) ρ[13] is the mathematical quantity that represents the quantum physical state. It is especially useful for describing mixtures, physical states that cannot be described by a vector or a wave function. A special case that is frequently discussed is a "pure state."

1.4.3 Pure State—Fundamental Postulate IIa

The pure state of a physical system, which is associated with an ensemble of microphysical objects, is described by a vector ϕ (up to a phase) or, equivalently, by a state operator $\rho_\phi = |\phi\rangle\langle\phi|$ that projects onto a one-dimensional space. In this state

---

[13] The matrix elements $\langle E_n|\rho|E_n'\rangle$ form the density matrix.

the expectation value of an observable A is given by

$$\sum_n \frac{N_n}{N} a_n \doteqdot \mathrm{Tr}(A|\phi\rangle\langle\phi|) = \langle\phi|A|\phi\rangle\,, \tag{1.4.6}$$

which is the matrix element of A between the state vector ϕ. If a state is prepared in such a way, that the measurement of the observable A always results in one value a, then ϕ is an eigenvector of A with the eigenvalue a.

Because no measurement in physics is absolutely precise, agreement between the measured and predicted quantities is expected only to within a certain error. As opposed to classical physics, in quantum physics an entirely new feature comes into play: theoretical quantities such as $\mathrm{Tr}(A\rho)$ and, in the case of a pure state $\langle\phi|A|\phi\rangle$ or $\langle\phi|\Lambda_n|\phi\rangle = \langle\phi|n\rangle\langle n|\phi\rangle$ represent the average values of measurements and do not predict the result of a single measurement, which in general is predicted only to within a theoretical error. In quantum theory there are theoretical uncertainties.

1.4.4 Uncertainties and Standard Deviations

In the energy-loss experiment, the energy E_0 is obtained N_0 times, the energy E_1 is obtained N_1 times, ..., and the energy E_7 is obtained N_7 times. There are a total of N measurements where $N = N_0 + N_1 + \cdots + N_7$. For the N measurements, the average value obtained for energy is denoted $\langle E\rangle$ and is given by (1.3.11). The statistical deviation from the average value $\langle E\rangle$ is denoted ΔE and is given by the root-mean-square deviation (or standard deviation) defined by

$$\Delta E = \sqrt{\sum_{n=0}^{7} \frac{N_n}{N}(E_n - \langle E\rangle)^2} = \sqrt{\sum_{n=0}^{7} \frac{N_n E_n^2}{N} - \langle E\rangle^2} = \sqrt{\langle E^2\rangle - \langle E\rangle^2}\,. \tag{1.4.7}$$

Generalizing to an arbitrary linear operator $A = P,\ Q,\ PQ$ etc, for which the value a_i is obtained N_i times, the root-mean-square deviation Δa from the average value $\langle a\rangle$ is

$$\Delta a = \sqrt{\sum_{i=0}^{\infty} \frac{N_i a_i^2}{N} - \langle a\rangle^2} = \sqrt{\langle a^2\rangle - \langle a\rangle^2}\,. \tag{1.4.8}$$

Using the notion of the "expectation value" of an observable A, the following statistical quantity is introduced: The *uncertainty* of an observable A in the pure state ϕ is defined by

$$\Delta_\phi A = \sqrt{\langle\phi|A^2|\phi\rangle - \langle\phi|A|\phi\rangle^2}\,. \tag{1.4.9}$$

Using the fact that the statistical operator for a pure state ϕ is $\rho = \Lambda_\phi = |\phi\rangle\langle\phi|$, the above equation can be rewritten as (Problem 1.17)

$$\Delta_\phi A = \sqrt{\mathrm{Tr}(\Lambda_\phi A^2) - (\mathrm{Tr}\,\Lambda_\phi A)^2}\,. \tag{1.4.10}$$

For a system in a state characterized by the general statistical operator ρ, *the uncertainty of an observable A in a general state ρ*, $\Delta_\rho A$ is defined by a straightforward generalization of (1.4.10):

$$\Delta_\rho A = \sqrt{\mathrm{Tr}(\rho A^2) - (\mathrm{Tr}(\rho A))^2} = \sqrt{\langle A^2\rangle_\rho - (\langle A\rangle_\rho)^2}\,. \tag{1.4.11}$$

To obtain a more intuitive grasp of the meaning of the uncertainty of an observable, the uncertainty of the energy operator is calculated for the CO molecule in the state ρ given by (1.3.12) for the energy-loss experiment depicted in Fig. 1.7 on page 27. Inserting ρ of (1.3.12) into (1.4.11) for the observable H,

$$\begin{aligned}\Delta_\rho H &= \sqrt{\mathrm{Tr}(\rho H^2) - (\mathrm{Tr}(\rho H))^2} = \sqrt{\langle H^2\rangle_\rho - (\langle H\rangle_\rho)^2}\\ &= \sqrt{\mathrm{Tr}(\sum_{n=0}^{7} w_n \Lambda_n H^2) - (\mathrm{Tr}(\sum_{n=0}^{7} w_n \Lambda_n H))^2}\\ &= \sqrt{\sum_{n=0}^{7} w_n \langle E_n|H^2|E_n\rangle - (\sum_{n=0}^{7} w_n \langle E_n|H|E_n\rangle)^2}\\ &= \sqrt{\sum_{n=0}^{7} w_n E_n^2 - (\sum_{n=0}^{7} w_n E_n)^2}\,.\end{aligned} \tag{1.4.12}$$

The expression on the right-hand side of (1.4.12) is always non-negative, even when the number of terms in the sum is infinite instead of just seven. This result follows from the Cauchy inequality for infinite sequences:

$$(\sum_n^\infty x_n y_n)^2 \le (\sum_n^\infty x_n^2)(\sum_n^\infty y_n^2)\,. \tag{1.4.13}$$

Taking $x_n = \sqrt{w_n}$ and $y_n = E_n\sqrt{w_n}$, the above equality takes the desired form,

$$(\sum_n^\infty w_n E_n)^2 \le (\sum_n^\infty w_n)(\sum_n^\infty w_n E_n^2) = \sum_n^\infty w_n E_n^2 \,. \tag{1.4.14}$$

The final equality above follows from (1.3.9).[14]

Equation (1.4.7) gives the statistical deviation ΔE from the average value $\langle E \rangle$ of the measurement (1.3.11) for the energy. The experimental deviation ΔE is associated with the theoretical uncertainty of the energy operator $\Delta_\rho H$ of quantum theory as given in (1.4.12),

$$\Delta_\rho H \doteqdot \Delta E \,. \tag{1.4.15}$$

In general the quantum theoretical uncertainty $\Delta_\rho A$ of the observable A in (1.4.11) is the theoretical analogue of the experimental *root-mean-square deviation* Δa in (1.4.8).

$$\Delta_\rho A \doteqdot \Delta a \,. \tag{1.4.16}$$

The uncertainty $\Delta_\rho H$ is equal to zero if $w_n = \delta_{nk}$, which is true only if $\rho = \Lambda_{|E_k\rangle} = |E_k\rangle\langle E_k|$. This means that the uncertainty of the energy H in the state ρ, denoted $\Delta_\rho H$, is equal to zero only if ρ is a pure energy eigenstate $\Lambda_{|E_k\rangle}$. For instance, if the system is not in the mixed state ρ of (1.3.12) but is instead in a pure energy eigenstate $|E_n\rangle\langle E_n|$, then the uncertainty for the energy operator is

$$\Delta_{|E_n\rangle} H = \sqrt{\langle H^2\rangle_{|E_n\rangle} - \langle H\rangle_{|E_n\rangle}^2} = \sqrt{\langle E_n|H^2|E_n\rangle - \langle E_n|H|E_n\rangle^2} = \sqrt{E_n^2 - E_n^2} = 0 \,. \tag{1.4.17}$$

The value of "uncertainty" $\Delta_\rho H$ is thus a theoretical expression for the uncertainty in the outcome of the single measurement of the observable H.

If the system is in the pure state $|E_n\rangle$, all molecules of the ensemble described by $\Lambda_{|E_n\rangle} = |E_n\rangle\langle E_n|$ should be "the same"; therefore, each molecule should have the same energy E_n, each measurement should lead to the same value, and the uncertainty $\Delta_{|E_n\rangle} H$ for the energy H should be zero. The probability $w_n = 1$, and the probabilities for all other energy values should be zero, $w_p = 0$ for $E_p \neq E_n$. The uncertainty deviates from zero if more than one value w_n is different from zero (Problem 1.18). The larger the number of non-zero w_n, the larger the uncertainty $\Delta_\rho H$, where the w_n satisfy the condition (1.3.9). If the system is in a mixture $\rho = \sum w_n |E_n\rangle\langle E_n|$ with $w_0, w_1, \ldots, w_7$ all different from zero, then the molecules

---

[14] As shown in Example 1.3.1 on page 32, the condition $\mathrm{Tr}(\rho) = 1$, which is part of Fundamental Postulates II, implies (1.3.9).

are in a mixture of eight kinds of molecules, one with energy E_0, the next with E_1, etc. In a measurement of H, each value E_n is obtained with probability $w_n \neq 1$, and the uncertainty of H, as given in (1.4.12), is different from zero. For a mixture of molecules with different energy values, it is expected that the root mean square deviation (1.4.12) from the average energy value $\langle E\rangle = \langle H\rangle_\rho$ is non-zero since the mixture contains molecules with different energies E_n.

If the CO molecules were all in a single energy eigenstate $|E_k\rangle$, then a standard deviation of zero is expected: at least in classical physics all CO molecules are “the same” and, therefore, all values of their observables should be the same. This is indeed the case for the energy observable H in the energy eigenstate

$$|E_k\rangle : \Delta_{\Lambda_k} = \sqrt{\mathrm{Tr}(\Lambda_{E_k} H^2) - (\mathrm{Tr}(\Lambda_k H))^2} = \sqrt{E_k^2 - E_k^2} = 0\,.$$

But for the expectation values of other operators such as Q, Q^2, P, and P^2, in an energy eigenstate, this is not the case. To calculate the expectation values $\langle Q\rangle$, $\langle Q^2\rangle$, $\langle P\rangle$, and $\langle P^2\rangle$ in the state ρ of CO molecules, the following matrix elements, which are calculated in Problem 1.6, are required:

$$\langle E_n|Q|E_n\rangle = 0 \qquad \langle E_n|P|E_n\rangle = 0 \tag{1.4.18a}$$

$$\langle E_n|Q^2|E_n\rangle = \frac{\hbar}{2\mu\omega}(2n+1) \qquad \langle E_n|P^2|E_n\rangle = \frac{\mu\omega\hbar}{2}(2n+1)\,. \tag{1.4.18b}$$

For the ensemble of CO molecules in the state ρ of (1.3.12),

$$\langle Q\rangle_\rho = \sum_{n=0}^{7} w_n \langle E_n|Q|E_n\rangle = 0\,, \tag{1.4.19a}$$

$$\langle P\rangle_\rho = \sum_{n=0}^{7} w_n \langle E_n|P|E_n\rangle = 0\,, \tag{1.4.19b}$$

$$\langle Q^2\rangle_\rho = \sum_{n=0}^{7} w_n \langle E_n|Q^2|E_n\rangle = \sum_{n=0}^{7} w_n \frac{\hbar}{2\mu\omega}(2n+1)\,, \tag{1.4.19c}$$

$$\langle P^2\rangle_\rho = \sum_{n=0}^{7} w_n \langle E_n|P^2|E_n\rangle = \sum_{n=0}^{7} w_n \frac{\mu\omega\hbar}{2}(2n+1)\,. \tag{1.4.19d}$$

Using (1.4.19) the uncertainty of the measurement of Q or P can be calculated using the definition (1.4.11) (Problem 1.19),

$$\Delta_\rho Q = \sqrt{\langle Q^2\rangle_\rho - (\langle Q\rangle_\rho)^2} = \sqrt{\sum_{n=0}^{7} w_n \frac{\hbar}{2\mu\omega}(2n+1)}\,, \tag{1.4.20a}$$

$$\Delta_\rho P = \sqrt{\langle P^2\rangle_\rho - (\langle P\rangle_\rho)^2} = \sqrt{\sum_{n=0}^{7} w_n \frac{\mu\omega\hbar}{2}(2n+1)}\,, \tag{1.4.20b}$$

and find that in the mixed state ρ of (1.3.12), the uncertainties of P and Q are different from zero. This is not a surprise since ρ is a mixture of several energy eigenstates that contain molecules with different energies E_n, $n = 0, 1, \ldots$.

The uncertainties are now calculated for the operators Q and P in a *pure state* $\Lambda_{|E_k\rangle} = |E_k\rangle\langle E_k|$ with a definite energy value E_k. Then all the molecules of this ensemble have the same energy, namely E_k, so they are completely identical. The uncertainties are calculated using (1.4.19) with the result

$$\Delta_{|E_k\rangle} Q = \sqrt{\langle Q^2\rangle_{|E_k\rangle} - (\langle Q\rangle_{|E_k\rangle})^2} = \sqrt{\frac{\hbar}{2\mu\omega}(2k+1)} \tag{1.4.21a}$$

$$\Delta_{|E_k\rangle} P = \sqrt{\langle P^2\rangle_{|E_k\rangle} - (\langle P\rangle_{|E_k\rangle})^2} = \sqrt{\frac{\mu\omega\hbar}{2}(2k+1)} \tag{1.4.21b}$$

In a pure state in quantum mechanics, the uncertainty for an observable is, in general, different from zero!

Even under ideal conditions, when the values of the position operator Q are measured in a pure energy eigenstate $\rho = \Lambda_{|E_k\rangle}$, different measurements yield different values. For instance, the values x_1 is obtained with a frequency N_1/N, x_2 is obtained with a frequency $N_2/N \ldots$, where $N = \sum_i N_i$. The average value is

$$\langle x\rangle = \sum_i \frac{N_i}{N} x_i\,. \tag{1.4.22}$$

This average value is calculated in quantum theory as the expectation value of the position operator in the energy eigenstate $|E_k\rangle$. According to (1.4.19a),

$$\langle Q\rangle_{|E_k\rangle} = \langle E_k|Q|E_k\rangle = \mathrm{Tr}(Q|E_k\rangle\langle E_k|) = 0\,. \tag{1.4.23}$$

In an experiment with a large number of measurements, the two values should agree,

$$\langle Q\rangle_{|E_k\rangle} \doteqdot \langle x\rangle\,. \tag{1.4.24}$$

The theoretical error in the expected value of Q is calculated in (1.4.20a). Since $\varDelta_{|E_k\rangle} Q \neq 0$, even in the ideal situation where all oscillators are in the pure energy eigenstate $\varLambda_{|E_k\rangle} = |E_k\rangle\langle E_k|$, the values of the position operator Q are only determined to within a quantum theoretical error $\varDelta_{|E_k\rangle} Q$. The quantum mechanical uncertainty $\varDelta_{|E_k\rangle} Q$ agrees with the root-mean-square deviation $\varDelta x$ in an experiment with a large number of measurements:

$$\varDelta_{|E_k\rangle} Q \doteqdot \varDelta x = \sqrt{\sum_i \frac{N_i}{N} x_i^2 - \langle x\rangle^2}\,. \tag{1.4.25}$$

Although the expectation values (1.4.19a) are zero for every energy eigenstate $|E_k\rangle$, the predicted root-mean-square deviations (1.4.20a) are all different from zero.

Summarizing, the quantum theoretical expectation value (1.4.3) of the observable A in the state ρ corresponds to the average value of the measurement (1.4.4), and the quantum theoretical uncertainty (1.4.11) corresponds to the root-mean-square deviation from the "measured" value (1.4.8),

$$\langle A\rangle_\rho \doteqdot \langle a\rangle, \quad \varDelta_\rho A \doteqdot \varDelta a\,. \tag{1.4.26}$$

Specifically, the expectation value of the observable H in (1.3.15a) that is in the mixed state ρ of (1.3.12) for the energy-loss experiment Fig. 1.7 on page 27 corresponds to the measured average value as expressed by (1.3.11). And the uncertainty of the observable H in the mixed state ρ of (1.3.12) as given by (1.4.12) corresponds to the root mean square deviation (1.4.7).

1.4.5 Heisenberg Uncertainty Principle

Even in pure states the outcome of the measurement of an observable is, in general, uncertain. Statistics cannot in principle be eliminated from quantum mechanics; systems in the "same pure state" do not give the same value in identical measurements. One measurement of an observable A gives the value a_1 and a second, identical measurement can give the value a_2, etc. All that can be determined theoretically is the probability w_n that a certain value $a_n = \langle E_n|A|E_n\rangle$ can be expected in the measurement of an observable A.

For the k-th energy eigenstate of the harmonic oscillator, these "theoretical root-mean-square deviations" called "uncertainties" of the position and momentum operator are given by (1.4.21). Using (1.4.21) to calculate the product $\varDelta_{|E_k\rangle} P\, \varDelta_{|E_k\rangle} Q$ in the k-th energy eigenstate,

$$\varDelta_{|E_k\rangle} P\, \varDelta_{|E_k\rangle} Q = \hbar(k + \frac{1}{2})\,.$$

For $k = 0$ the above equation becomes,

$$\Delta_{|E_0\rangle} P \, \Delta_{|E_0\rangle} Q = \frac{\hbar}{2} \, .$$

Thus for any pure energy eigenstate of the harmonic oscillator

$$\Delta P \, \Delta Q \geq \frac{\hbar}{2} \, . \tag{1.4.27}$$

This above relation is true in general and is called the Heisenberg uncertainty relation. It reveals that it is impossible to simultaneously know both the position and momentum with arbitrary precision. The Heisenberg uncertainty relation is a mathematical consequence[15] of the Heisenberg commutation relations $[Q, P] = i\hbar\mathbb{1}$.

1.5 Radiative Transitions

1.5.1 Dipole Radiation

In the energy-loss experiment depicted in Fig. 1.7 on page 27, the energy levels of the vibrating CO molecule were measured by exploiting energy conservation in the process (1.3.2). Specifically, the energy ΔE of (1.4.3) lost by an electron equals the energy gained by the CO molecule. The difference between adjacent energy levels of the CO molecule are determined by measuring the energy lost by electrons as a function of sweep voltage as depicted in Fig. 1.8 on page 28. The electrons loose only discrete amounts of energy as shown by the equally spaced peaks, indicating that the energy levels of the CO molecule are also equally spaced 0.265 eV apart.

The most significant structure of a vibrating diatomic molecule—CO or any molecule consisting of unlike atoms—is its electric dipole moment. The *dynamical* picture of the CO molecule is a carbon and an oxygen nucleus with electrons swirling around them. The centers of positive and negative charge do not coincide for molecules consisting of unlike atoms, and the electric dipole moment is the vector directed from the center of negative charge to the center of positive charge given by

$$D_0^{\text{classical}} = qd \, , \tag{1.5.1}$$

where q is the charge and d is the distance between the centers of the positive and negative charge.

---

[15] A. Bohm, *Foundations and Applications* 2nd ed., Springer-Verlag, Berlin, Heidelberg, New York, 1986.

When an electric dipole oscillates, the charges are accelerated so they emit electromagnetic radiation. In quantum mechanics, a photon is emitted when a transition is made from the excited oscillator state O_m of the CO molecule to a lower oscillator state O_{m-1},

$$O_m \to O_{m-1} + \gamma \,, \tag{1.5.2}$$

Energy conservation for this process yields

$$E_m = E_{m-1} + E_\gamma \,. \tag{1.5.3}$$

Using the fact that the energy of the photon is $E_\gamma = h\nu_\gamma$, where ν_γ is the frequency of the emitted electromagnetic radiation, the above equation becomes

$$E_m - E_{m-1} = h\nu_\gamma = \hbar\omega_\gamma \,. \tag{1.5.4}$$

The energy difference on the left-hand side of (1.5.8) can be calculated from (1.2.40) for the energy levels of the oscillator,

$$E_\gamma = E_m - E_{m-1} = \hbar\omega\left(m + \frac{1}{2}\right) - \hbar\omega\left(m - 1 + \frac{1}{2}\right) = \hbar\omega \,. \tag{1.5.5}$$

From (1.5.5) it follows that the emitted light has an angular frequency ω_γ that equals the angular frequency $\omega = \sqrt{k/m}$ of the vibrating molecule. Here k is the spring constant (1.2.2) of the diatomic vibrator and μ is its reduced mass (1.2.6). That is, the electromagnetic radiation is emitted at the characteristic frequency of the vibrating diatomic molecules.

When the centers of charge $(-q)$ and $(+q)$ are not vibrating, the permanent dipole moment $D_0^{\text{classical}}$ of the molecule lies along the internuclear axis. If the interatomic (or internuclear) distance changes, the dipole moment will change. To a good approximation the dipole moment is a linear function of the deviation from the equilibrium position d of the interatomic distance,

$$D^{\text{classical}} = q(d + x) = D_0^{\text{classical}} + q\,x \,. \tag{1.5.6}$$

The change in position x of the centers of positive and negative charge oscillates with the frequency of the mechanical vibration, so the electric dipole moment oscillates with the same frequency.

From the correspondence principle, the transition from the classical to the quantum oscillator is accomplished by making the replacement $x \to Q$ with the result that the classical dipole moment $D_0^{\text{classical}} = qd$ in (1.5.6) is replaced by the dipole operator

$$D = D_0 + q\,Q \,, \tag{1.5.7}$$

where Q is the operator of the relative position of the center of positive charge $(+q)$ and negative charge $(-q)$. The permanent dipole moment D_0 of the molecule does not act as an operator in the space of oscillator states but will play a role when rotation of the diatomic molecule in the three- dimensional space is discussed in the next chapter. Thus the dipole operator for radiation between the oscillator states is the operator

$$\mathbf{D} = q\mathbf{Q}\,. \tag{1.5.8}$$

For molecules consisting of two like atoms such as O_2 or N_2, the electric dipole moment operator (1.5.6) is the zero operator because the centers of positive and negative charge coincide. Oscillations of such molecules about the equilibrium position do not lead to oscillations of the centers of charge, so no emission or absorption of electromagnetic radiation occurs, and there are no transitions between different energy levels

The average power $\overline{R}_{nm}$[16] emitted by dipole radiation for transitions from the oscillator energy eigenstates $|E_m\rangle$ to the states $|E_n\rangle$[17] of the quantum system is given by

$$\overline{R}_{nm} = \frac{\omega_\gamma^4\, q^2}{3\pi\epsilon_0 c^3}|\langle E_n|Q|E_m\rangle|^2\,. \tag{1.5.9}$$

Since the matrix element $\langle E_n|Q|E_m\rangle$ as given by (1.2.52) is zero unless $n = m \pm 1$, the average power $\overline{R}_{nm}$ radiated by a dipole is zero unless

$$n = m - 1 \quad \text{or} \quad n = m + 1\,. \tag{1.5.10}$$

The above rules are called the selection rules for dipole transitions.

Because the interaction between the electron and diatomic oscillators in the energy-loss experiment (1.3.2) is entirely different from the dipole interactions of a CO molecule with the electromagnetic field, the energy-loss experiment is not subject to the restrictions given by the selection rules (1.5.10), and transitions to all excited states are possible as long as energy conservation is fulfilled (and as long as the CO molecule remains an oscillator).

---

[16] This formula is derived in many quantum mechanics texts. See, for example, D. Park, *Introduction to the Quantum Theory*, 3rd Ed., McGraw-Hill, New York, 1992.

[17] In quantum mechanics it is conventional to label quantities first by the quantum number(s) of the final state—in this case by n—and then by the quantum number(s) of the initial state—in this case by m.

For the process (1.5.2), energy conservation (1.5.3) allows emission to occur only when the final state, labeled by n, satisfies $n = m - 1$. Dipole transitions can also occur when a dipole absorbs a photon in the process

$$\gamma + O_m \to O_{m+1}\,. \tag{1.5.11}$$

Energy conservation for the above process yields

$$E_\gamma + E_m = E_{m+1}\,. \tag{1.5.12}$$

Energy conservation requires that the final state of the oscillator, characterized my the number n satisfy $n = m + 1$. Using the same logic that lead to (1.5.5),

$$E_\gamma = E_{m+1} - E_m = \hbar\omega\left(m + 1 + \frac{1}{2}\right) - \hbar\omega\left(m + \frac{1}{2}\right) = \hbar\omega = h\nu_\gamma = \hbar\omega_\gamma\,. \tag{1.5.13}$$

From (1.5.5) and (1.5.13) it follows for either photon emission as depicted in (1.5.2) or for photon absorption shown in (1.5.11), the angular frequency of the emitted or absorbed photon equals the characteristic angular frequency of the oscillator, $\omega_{\text{oscillator}} = \sqrt{k/\mu} = \omega_\gamma$. Thus the frequency of radiated light is independent of the energy level n. In summary, for the *particular* case of the quantum mechanical harmonic oscillator, the frequency of emitted and absorbed light is the same from all energy levels and is equal to the frequency of the vibrating diatomic molecule.

Other electromagnetic transitions such as quadrupole transitions due to the matrix elements $\langle n|Q \cdot Q|m\rangle$ and higher multipole transitions are possible; however, the contributions of these matrix elements are smaller by the ratio[18] of the Bohr radius r_B to the wavelength, $r_B/\lambda \sim 0.5 \times 10^{-8}\text{cm}/0.5 \times 10^{-4}\text{cm} \sim 10^{-4}$. Since the intensity is proportional to the square of the magnitude of the matrix element, the intensity is smaller by a factor of 10^{-8} and is too small to be observed.

1.5.2 Einstein Coefficients

Beginning with the formula (1.5.9) for the average power $\overline{R}_{nm}$ emitted during the spontaneous emission of a photon in the process (1.5.2), it is possible to calculate the transition rate, the number of oscillators that make the transition $O_m \to O_n + \gamma$ per unit time. Since the observed quantum mechanical quantity is a probability, the transition rate for dipole radiation between neighboring energy levels of the oscillator must also be expressible in terms of probabilities. The probability is

---

[18] C. Cohen-Tannoudji, B. Diu, F. Laloë, *Quantum Mechanics* v. 2, p. 1313, Wiley &Sons, New York, NY, 1977.

expressed in terms of the square of the absolute value of the matrix element $|\langle E_n|Q|E_m\rangle|^2$ in (1.5.9). The quantity $|\langle E_n|Q|E_m\rangle|^2$ is the probability of observing the property $|E_n\rangle\langle E_n|$ in the state $\phi = Q|E_m\rangle$.

$$\begin{pmatrix}\text{Probability for tran-}\\ \text{sition from the state}\\ \phi \text{ to the state } |E_n\rangle\end{pmatrix} = \mathrm{Tr}(|E_n\rangle\langle E_n|\phi\rangle\langle\phi|) = |\langle E_n|\phi\rangle|^2 = |\langle E_n|Q|E_m\rangle|^2. \tag{1.5.14}$$

The quantity $\langle E_n|\phi\rangle$ is, therefore, called the transition probability amplitude or just the transition amplitude for a transition from the state ϕ to the energy eigenstate $|E_n\rangle$. In the same spirit $\langle E_n|Q|E_m\rangle$ is called the transition amplitude for a transition performed by the observable Q between the energy eigenstates $|E_n\rangle$ and $|E_m\rangle$.

According to the statistical interpretation of quantum mechanics, the ensemble of CO molecules has N_m oscillators in the energy eigenstate $|E_m\rangle$. It is not possible to say which of the oscillators will make the transition to the state $|E_n\rangle$ because the transitions occur randomly. After a time Δt the number of oscillators that have made the transition to the state $|E_n\rangle$ is denoted $N_{nm} < N_m$. The time Δt must be large enough so that N_{nm} is sufficiently large that a repetition of the experiment will yield approximately the same fraction of transitions N_{nm}/N_m. The probability for the emission of a photon with energy $\hbar\omega_{nm} = E_m - E_n$ during the time Δt is then N_{nm}/N_m.

Since a transition from the energy eigenstate $|E_m\rangle$ to the eigenstate $|E_n\rangle$ leads to the emission of the energy $E_m - E_n = \hbar\omega_{nm}$, the average energy emitted by the ensemble is $(N_{nm}/N_m)\hbar\omega_{nm}$. This equals the average emitted power $\overline{R}_{nm}$ multiplied by the time interval Δt, $\overline{R}_{nm}\Delta t$. From the correspondence between the classical and quantum expressions we therefore obtain

$$\text{average emitted energy} \equiv \frac{N_{nm}}{N_m}\hbar\omega_{nm} = \overline{R}_{nm}\Delta t = \frac{\omega_\gamma^4\, q^2}{3\pi\epsilon_0 c^3}|\langle E_n|Q|E_m\rangle|^2\Delta t\,. \tag{1.5.15}$$

The final equality is obtained with the aid of the explicit expression (1.5.9) for $\overline{R}_{nm}$. The probability of a transition from $|E_m\rangle$ to $|E_n\rangle$ per unit time is

$$\begin{pmatrix}\text{probability of a transition from}\\ |E_m\rangle \text{ to } |E_n\rangle \text{ per unit time}\end{pmatrix} = \frac{N_{mn}}{N_m}\frac{1}{\Delta t}\,. \tag{1.5.16}$$

The desired expression for the transition probability per unit time is immediately obtained by solving (1.5.15) for $N_{mn}/(N_m\Delta t)$ and using $\omega_{mn} = \omega_\gamma$ that follows from (1.5.5),

$$\frac{N_{mn}}{N_m}\frac{1}{\Delta t} = \frac{\omega_\gamma^3\, q^2}{3\pi\epsilon_0 c^3\hbar}|\langle E_n|Q|E_m\rangle|^2\,. \tag{1.5.17}$$

The *transition probability per unit time* or the *transition rate* for a transition from level $|E_m\rangle$ to level $|E_n\rangle$ as a result of the spontaneous emission of photon with energy E_γ of (1.5.9) is denoted A_{mn}[19] and is called the Einstein coefficient[20] for spontaneous emission. According to (1.5.18),

$$A_{mn} = \frac{\omega_\gamma^3\, q^2}{3\pi\epsilon_0 c^3\hbar}|\langle E_n|Q|E_m\rangle|^2\,. \tag{1.5.18}$$

Since A_{mn} is the probability of decay per unit time, the lifetime τ_{mn} of the excited state $|E_m\rangle$ for decay into the state $|E_n\rangle$ is the reciprocal of A_{mn},

$$\tau_{nm} = \frac{1}{A_{mn}} = \frac{3\pi\epsilon_0 c^3\hbar}{\omega_\gamma^3\, q^2|\langle E_n|Q|E_m\rangle|^2}\,. \tag{1.5.19}$$

If an ensemble of diatomic molecules is placed inside a backbody radiator, in addition to spontaneous emission of photons, the molecules can undergo two additional electromagnetic transitions. In a black body at temperature T, the energy distribution of photons as a function of the photon angular frequency ω is[21]

$$u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{1}{e^{\hbar\omega/kT} - 1}\,, \tag{1.5.20}$$

where $u(\omega)\mathrm{d}\omega$ is the energy density of the photons in the angular frequency range ω to $\omega + \mathrm{d}\omega$ and k is Boltzmann's constant.

The three electromagnetic transitions that oscillators can undergo in a black body are as follows:

1. *Spontaneous Emission* has already been discussed: the oscillator O_m in the state $|E_m\rangle$ emits a photon in the transition $O_m \to O_n + \gamma$, where $n = m - 1$. The probability of a transition per unit time is the Einstein coefficient A_{mn} given in (1.5.18). In the black body if there are N_m oscillators in the state $|E_m\rangle$, the total number of oscillators that make the transition from state $|E_m\rangle$ to state $|E_n\rangle$ per unit time is

$$\begin{pmatrix}\text{number of oscillators undergoing}\\ \text{spontaneous emission per unit}\\ \text{time}\end{pmatrix} = N_m A_{mn}\,. \tag{1.5.21}$$

---

[19] Using Einstein's notation, the Einstein coefficients are labeled first by the quantum number(s) of the initial state—in this case by m—and then by the quantum number(s) of the final state—in this case by n.

[20] A. Einstein, Physik. Z.,**18**, 121(1917).

[21] The photon energy density $u(\omega)$ or the photon angular frequency density is the time average of the magnitude of the Poynting vector divided by c.

2. *Induced Emission* is the process by which an oscillator O_m is induced by the presence of electromagnetic radiation to emit an electron in the transition $O_m \rightarrow O_n + \gamma$. The angular frequency of the radiation must be at or near the angular frequency $\omega = (E_m - E_n)/\hbar$ of the emitted photon, just as the driving frequency must be near the resonant frequency of a system to get a large response. The probability of induced emission per unit time is proportional to the density of radiation at the angular frequency $\omega = (E_m - E_n)/\hbar$. The proportionality constant is defined to be the Einstein coefficient B_{mn} so the probability of a transition per unit time is $u(\omega)B_{mn}$. In the black body if there are N_m oscillators in the state $|E_m\rangle$, the total number of oscillators that are induced to make the transition from state $|E_m\rangle$ to state $|E_n\rangle$ per unit time is

$$\begin{pmatrix}\text{number of oscillators undergoing}\\ \text{induced emission per unit time}\end{pmatrix} = N_m u(\omega) B_{mn}\,. \tag{1.5.22}$$

3. *Spontaneous Absorption* is the process by which an oscillator O_n in the initial state $|E_n\rangle$ absorbs a photon in the transition $O_n + \gamma \rightarrow O_m$, where $m = n + 1$. The probability of a transition per unit time is proportional to the density of radiation at the frequency $\omega = (E_m - E_n)/\hbar$. The proportionality constant is defined to be the Einstein coefficient B_{nm} so the probability of a transition per unit time is $u(\omega)B_{nm}$. In the black body if there are N_n oscillators in the state $|E_n\rangle$, the total number of oscillators that are induced to make the transition from state $|E_n\rangle$ to state $|E_m\rangle$ per unit time is

$$\begin{pmatrix}\text{number of oscillators undergoing}\\ \text{induced absorption per unit time}\end{pmatrix} = N_n u(\omega) B_{nm}\,. \tag{1.5.23}$$

In thermal equilibrium the total number of oscillators emitting photons per unit time must equal the total number of oscillators absorbing photons per unit time:

$$\begin{pmatrix}\text{number of}\\ \text{oscillators}\\ \text{undergoing}\\ \text{spontaneous}\\ \text{emission per unit}\\ \text{time}\end{pmatrix} + \begin{pmatrix}\text{number of oscil-}\\ \text{lators undergoing}\\ \text{induced emission}\\ \text{per unit time}\end{pmatrix} = \begin{pmatrix}\text{number of oscil-}\\ \text{lators undergoing}\\ \text{induced absorption}\\ \text{per unit time}\end{pmatrix}. \tag{1.5.24}$$

Using (1.5.21)–(1.5.23) the above equation can be written as

$$N_m A_{mn} + N_m u(\omega) B_{mn} = N_n u(\omega) B_{nm}\,. \tag{1.5.25}$$

Solving for $u(\omega)$,

$$u(\omega) = \frac{A_{mn}}{\frac{N_n}{N_m} B_{nm} - B_{mn}}\,. \tag{1.5.26}$$

At thermal equilibrium Gibbs distribution[22] is

$$\frac{N_n}{N_m} = \frac{e^{-E_n/kT}}{e^{-E_m/kT}} = e^{(E_m - E_n)/kT} = e^{\hbar\omega/kT} , \tag{1.5.27}$$

where (1.5.5) has been used to obtain the final equality. Substituting (1.5.27) into (1.5.26), the expression for $u(\omega)$ becomes

$$u(\omega) = \frac{A_{mn}}{e^{\hbar\omega/kT} B_{nm} - B_{mn}} . \tag{1.5.28}$$

The above expression for $u(\omega)$ agrees with (1.5.20) iff

$$B_{nm} = B_{mn} \quad \text{and} \quad \frac{A_{mn}}{B_{mn}} = \frac{\hbar\omega^3}{\pi^2 c^3} . \tag{1.5.29}$$

From (1.5.18) it then immediately follows that Einstein's coefficient B_{mn} is given by

$$B_{mn} = \frac{\pi q^2}{3\epsilon_0 \hbar^2} |\langle E_n|Q|E_m\rangle|^2 . \tag{1.5.30}$$

1.5.3 Comparison with Experimental Results

The theoretical results are now compared with experimental data. The frequency of emitted photons in dipole decay of CO diatomic molecules can be determined from the energy-loss spectrum of the CO molecules in Fig. 1.8 on page 28. From the voltage difference between the equally spaced peaks in the energy-loss spectrum, it follows that the energy difference between adjacent energy levels of the vibrating CO molecule is

$$\Delta E = 0.265\,\text{eV} . \tag{1.5.31}$$

The frequency of the emitted photon can be calculated using energy conservation (1.5.5),

$$\nu_\gamma = \frac{\Delta E}{h} = \frac{0.265\,\text{eV}}{4.14 \times 10^{-15}\,\text{eV s}} = 6.41 \times 10^{13}\,\text{Hz} . \tag{1.5.32}$$

---

[22] L. D. Landau and E. M. Lifshitz, *Statistical Physics 5* V.1 (3 ed.) Pergamon Press, Oxford, 1980).

The wavelength corresponding to this frequency is

$$\lambda = \frac{c}{\nu} = \frac{3.00 \times 10^8\,\mathrm{m/s}}{6.41 \times 10^{13}/\mathrm{s}} = 4.68 \times 10^{-6}\,\mathrm{m}\,. \tag{1.5.33}$$

In molecular spectroscopy it is customary to express the frequency ν as the wave number ν/c, where c is the speed of light in cm/s, although the wave number $\nu/c = 1/\lambda$ is actually the number of waves per cm. Instead of introducing a new symbol, the wave number is also denoted ν. The frequency ν and wave number ν are distinguished by the units in which they are measured. The wave number ν/c of the radiation emitted by the transition between the vibrational levels of CO is then

$$\nu = \frac{6.41 \times 10^{13}\,\mathrm{s}^{-1}}{3.00 \times 10^{10}\,\mathrm{cm/s}} = 2140\,\mathrm{cm}^{-1}\,. \tag{1.5.34}$$

From the above discussion it follows that vibrating CO molecules emit or absorb electromagnetic radiation only with the *single* frequency given by (1.5.34). That is, only one spectral line is detected in the near infrared region. To show where infrared radiation is in the electromagnetic spectrum and to give orders of magnitude for the energy, frequency, wave number, and wavelength of the various types of radiation involved in molecular spectroscopy, a table of the electromagnetic spectrum is shown in Fig. 1.9 on the next page.

In the experimental absorption or emission spectra for CO, there is only the single frequency at $\nu = 2140\,\mathrm{cm}^{-1}$, so the theoretical prediction is fulfilled to a high degree of accuracy. If the absorption spectrum is obtained with a thin layer of absorbing gas, then only a single, broad, intense absorption line (or band) is detected in the near infrared region with a wave length around $\lambda = 4.68 \times 10^{-6}$ m. For other diatomic molecules consisting of unlike atoms, similar data is obtained: for HCl the band lies at $\lambda = 3.46 \times 10^{-6}$ m. Also, bands do not appear for molecules consisting of like atoms such as O_2, N_2, and H_2 because, as previously discussed, the electric dipole moment is zero for diatomic molecules consisting of two identical atoms.

If the absorption is observed with thicker layers of gas, the intensity of absorption of the fundamental band naturally increases because more molecules are undergoing dipole transitions. In addition, a second band similar in form but much weaker than the first appears at approximately half the wavelength or double the frequency. If the number of molecules undergoing dipole transitions is increased still further by increasing the pressure of the gas, a third and possibly even a fourth and a fifth band appear with wavelengths that are approximately a third, a fourth and a fifth, respectively, of that of the first band. The frequencies of these bands are, respectively, approximately three, four and five times that of the first band as can be seen in Fig. 1.5 on page 23, which schematically shows the complete infrared spectrum of the HCl diatomic molecule. In this figure the lengths of the vertical lines that represent the bands give an indication of their intensity. However, the actual decrease in intensity is five times as great as that shown.

THE ELECTROMAGNETIC SPECTRUM

Energy E (eV)	Frequency ν (s^{-1})	Transitions	Radiation	Wave Number ν (cm^{-1})	Wavelength λ (cm)
		Nuclear Magnetic Resonance	Radio Waves		
5×10^{-6}	1.2×10^{9}	Spin Orientations in Magnetic Field		4×10^{-2}	25
		Electron Spin Resonance	Microwaves (radar)		
3.1×10^{-3}	7.5×10^{11}	Molecular Rotations		25	4×10^{-2}
5×10^{-2}	1.2×10^{13}			400	2.5×10^{-3}
		Molecular Vibrations	Infrared Region		
0.5	1.2×10^{14}			4×10^{3}	2.5×10^{-4}
1.55	3.8×10^{14}			1.25×10^{4}	8×10^{-5}
			Visible		
3.1	7.5×10^{14}			2.5×10^{4}	4×10^{-5}
		Valence Electronic Transitions			
6.2	1.5×10^{15}		Ultraviolet	5×10^{4}	2×10^{-5}
1.24×10^{3}	3×10^{17}			10^{7}	10^{-7}
		Inner Shell Electronic Transitions	X-rays		
1.24×10^{4}	3×10^{18}			10^{8}	10^{-8}
		Nuclear Transitions	Gamma Rays		

Fig. 1.9 Schematic diagram of the electromagnetic spectrum. Note that the scale is not linear. Boundaries between regions are somewhat arbitrary

The explanation for the presence of these additional bands with lower intensity is that the diatomic molecule is not quite a harmonic oscillator. In a *harmonic* oscillator the restoring force (1.2.2) increases indefinitely with increasing distance from equilibrium. However, it is clear that in the actual molecule, when the atoms are a great distance from one another, the attractive force decreases and ultimately approaches zero. The quantum-mechanical oscillator is an approximate model of a diatomic molecule that is valid only in a limited energy range. To describe the finer details of vibrating molecules at higher energies (higher incident energies of electrons in the energy-loss experiment), the anharmonic forces have to be taken into account. The energy levels of the anharmonic oscillator are not equidistant like those of the harmonic oscillator, but rather their separation decreases slowly with increasing n.

The energy levels and absorption spectrum for an anharmonic oscillator that deviates only slightly from a harmonic oscillator are shown in Fig. 1.10. For the sake of clarity, ΔE is shown to decrease faster than is actually observed in most cases. The selection rule (1.5.10), namely, $n - m = \pm 1$, holds only approximately for the anharmonic oscillator and applies only to the most intense transitions. Transitions with $n - m = \pm 2, \pm 3, \ldots$ can also appear, although with rapidly decreasing intensity. All these results can be calculated using perturbation theory, which will be discussed in Chap. 4.

The experimental facts are discussed here to demonstrate that simple, solvable quantum-mechanical models such as the harmonic oscillator, the Kepler model of the hydrogen atom, or the rigid rotor discussed in the next chapter describe only the principal structures of a microphysical system in nature and cannot be expected to describe all details. This is not a deficiency of a specific, simple model but is instead a general property of physical theories: models are idealizations and cannot be expected to reproduce experimental results with arbitrarily high precision. An

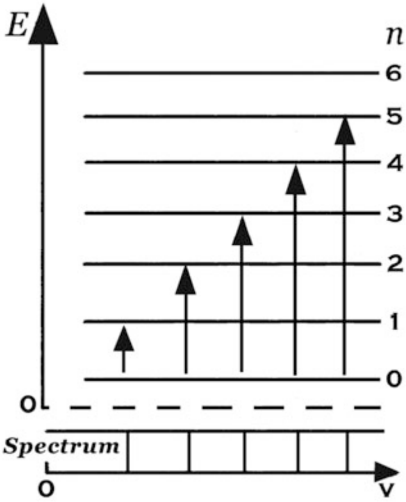


Fig. 1.10 Energy levels and infrared transitions of an anharmonic oscillator. The absorption spectrum is given schematically in the lower figure

actual physical system is understood by its agreement with *and* its deviation from these benchmarks. For a more detailed understanding, a finer benchmark is chosen. Examples of finer benchmarks are an anharmonic oscillator or the combination of two benchmarks such as a harmonic oscillator that also rotates. More detailed experimental results often reveal new features of a quantum system. The calculation of one additional decimal place in an experimental number often requires a new model and sometimes a completely new theory.

When a spectrometer with sufficiently high resolution is used to study the transition frequencies in the near infrared region in more detail, the $n = 1$ line for CO at about $\nu = 2140\ \text{cm}^{-1}$ is resolved into many individual, narrow lines as shown in Fig. 1.11. That is, instead of a single line around $\nu = 2140\ \text{cm}^{-1}$, there is instead a band called the vibration-rotation band. The experimental situation for HCl is similar: the $n = 1$ line of Fig. 1.5 on page 23 for HCl at about $\nu = 2890\ \text{cm}^{-1}$ is resolved into many individual, narrow lines as shown in Fig. 1.12. As can be seen from Fig. 1.11 for CO and Fig. 1.12 for HCl, these bands consist of a series of almost-equidistant lines with one line missing in the center of the band. Moving outward from the gap there are two branches: the P-branch is toward longer wavelengths, and the R-branch is toward smaller wavelengths.

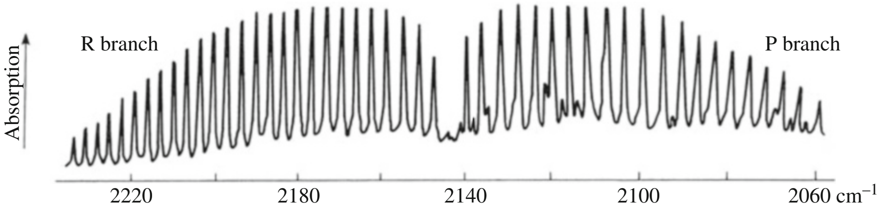


Fig. 1.11 The vibration-rotation band of carbon monoxide

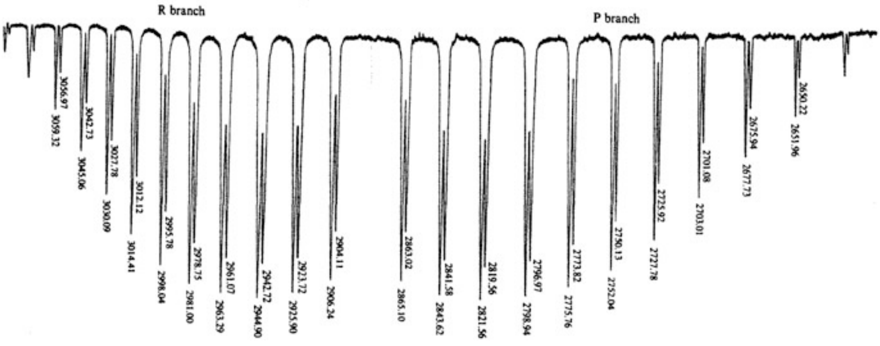


Fig. 1.12 The fundamental absorption band for HCl under high resolution. (The lines are doubled due to the presence of the two isotopes Cl^{35} and Cl^{37} in the ratio 3:1; this effect is not discussed here). [From N. L. Alpert, W. E. Keiser and H. A. Szymanski. *Theory and Practice of Infrared Spectroscopy*. Wiley, New York,1970. with permission]

Such fine structure in the absorption or emission spectra of electromagnetic radiation for the CO molecule would be expected if the energy levels of the vibrating molecule of Fig. 1.6 on page 25 were split into a series of sublevels as depicted in Fig. 1.13 on the following page, which shows the lowest energy levels of the energy spectrum of the vibrating molecule in Fig. 1.10 on page 53.

The description of such splittings cannot be explained solely by the oscillator model. Such splittings can only mean that a state characterized by the quantum number n is not a pure state. Instead it is a mixture of states with different energies characterized by n and at least one additional quantum number $\ell = 0, 1, 2, \ldots$. In the oscillator model the state characterized by n was a pure state described by a projection operator Λ_n onto a one-dimensional subspace spanned by $\Lambda_n \mathscr{H}$. The space of states of the diatomic molecule characterized by the quantum number n must have at least as many dimensions as energy levels. When the number of energy levels equals the dimension, then a one-dimensional subspace or a projection onto a one-dimensional subspace is associated with each value of energy. Consequently, the oscillator model alone describes only part of the properties of a diatomic molecule. To describe the finer details of the spectrum, the oscillator model must be combined with another model that describes these finer details and incorporates additional features of the diatomic molecule that have not yet been taken into account.

Classical physics suggests a possible explanation for the finer details in the spectra of diatomic molecules. The classical picture of the CO molecule is two atoms with masses m_1 and m_2 separated by a distance x. This classical object cannot only perform vibrations along the internuclear axis, but it can also perform rotations in three-dimensional space around its center of mass. As long as the diatomic molecule is in the vibrational ground state, that is, as long as the energy involved is less than 0.265 eV for CO, the molecule is a rigid rotator: it can be considered to be two point-like masses m_1 and m_2 fastened to the ends of a weightless, rigid rod of length x; therefore, the rigid-rotator model can be studied while ignoring possible oscillations. This analysis will be done in the following chapter, providing a description of the CO states that are characterized by the quantum number $n = 0$ and also an approximate description of each set of oscillator states with a given vibrational quantum number n. In Chap. 3 the oscillator and rotator models are combined to form the vibrating rotator or the rotating vibrator that provides a theoretical explanation for the detailed spectrum of the diatomic molecule depicted in Fig. 1.11 on the facing page and Fig. 1.12 on the preceding page.

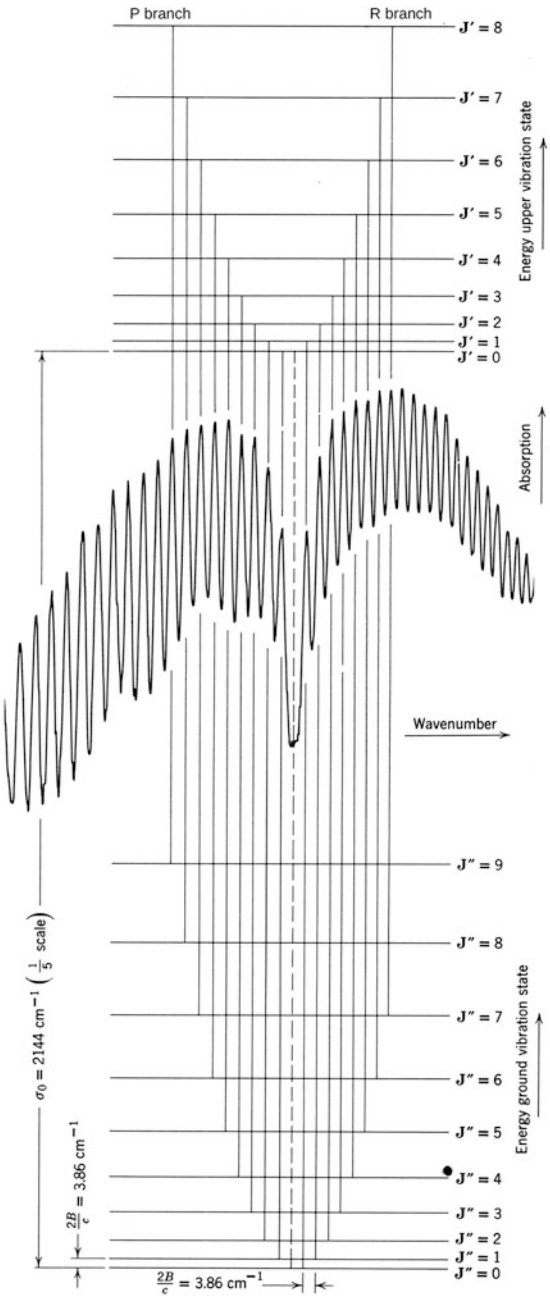


Fig. 1.13 Origin and appearance of rotational structure. P- and R-branches are shown to the left and right, respectively, on the spectrometer tracing of the CO fundamental absorption band at $2144\,\text{cm}^{-1}$. The Q-branch (dashed line) is missing. Energy levels are shown to scale, except that the distance between upper and lower vibrational states ($2144\,\text{cm}^{-1}$) should be about five times as great as in the figure. [From R. P. Bauman. *Absorption Spectroscopy*. Wiley, New York, 1962. with permission]

1.6 Summary

A classical harmonic oscillator consists of two mass points with respective masses m_1 and m_2 connected by an ideal spring that exerts a force proportional to the distance stretched from its equilibrium length. Ignoring the energy associated with the motion of the center of mass, the energy E of a classical harmonic oscillator is

$$E = \frac{1}{2}\mu\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \frac{1}{2}kx^2\,, \quad \frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}\,.$$

The quantum Hamiltonian is obtained from the classical expression for energy by replacing the classical expressions for momentum, $p = \mu\frac{\mathrm{d}x}{\mathrm{d}t}$, and position x, respectively, by the quantum mechanical momentum operator P and position operator Q:

$$H = \frac{P^2}{2\mu} + \frac{1}{2}kQ^2$$

Requiring that the momentum and position operators satisfy the Heisenberg commutation relations, the eigenvalues E_n of H are determined using algebraic techniques. Denoting a normalized eigenvector of H with an eigenvalue E_n by $|E_n\rangle$, where $H|E_n\rangle = E_n|E_n\rangle$, it follows that

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)\,, \quad \omega = \sqrt{\frac{k}{\mu}}\,, \quad n = 0, 1, 2, \ldots\,.$$

The (dimensionless) lowering and raising operators a and $a^\dagger$, respectively, are given by

$$a \equiv \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}Q + \frac{i}{\sqrt{\mu\omega\hbar}}P\right)\,, \qquad a^\dagger \equiv \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}Q - \frac{i}{\sqrt{\mu\omega\hbar}}P\right)\,,$$

and were introduced because they satisfy the equations

$$Ha|E_n\rangle = E_{n-1}\,a|E_n\rangle\,, \qquad Ha^\dagger|E_n\rangle = E_{n+1}\,a^\dagger|E_n\rangle\,.$$

From the above equations, if $a|E_n\rangle$ is unequal to zero, it is proportional to $|E_{n-1}\rangle$. Similarly, $a^\dagger|E_n\rangle$ is proportional to $|E_{n+1}\rangle$. Including proportionality factors,

$$a|E_n\rangle = \sqrt{n}\,|E_{n-1}\rangle\,, \quad a^\dagger|E_n\rangle = \sqrt{n+1}\,|E_{n+1}\rangle\,.$$

Using the raising operator $a^\dagger$, any normalized eigenvector $|E_n\rangle$ can be obtained from the eigenvector $|E_0\rangle$,

$$|E_n\rangle = \frac{1}{\sqrt{n!}}\,(a^\dagger)^n|E_0\rangle\ .$$

The vibrating diatomic molecule is one such physical system that can be approximately described as a quantum harmonic oscillator. From energy loss experiments, for diatomic molecules the energy levels are found to be equally spaced within a certain range of data. Thus, within this range, diatomic molecules can be described as quantum harmonic oscillators. For physical systems that can be described as harmonic oscillators, the dipole operator qQ is most often responsible for transitions between states.

An ensemble that cannot be further subdivided is called a pure state. It can be described either by a single vector $|E_n\rangle$ that spans the space or by the projection operator $\Lambda_n = |E_n\rangle\langle E_n|$ that projects onto the space. A mixture is an ensemble that can in principle be subdivided into pure states. If the mixture contains pure states $|E_n\rangle,\ n = 0, 1, 2, \ldots, N$, the mixture is described by the statistical or density operator ρ defined by

$$\rho = \sum_{n=0}^{N} w_n \Lambda_n = \sum_{n=0}^{N} w_n |E_n\rangle\langle E_n|\,,$$

where w_n is the probability of obtaining the value E_n. The expectation value of an operator A in the state characterized by ρ is

$$< A >_\rho = \mathrm{Tr}\,\rho A\ .$$

For the special case where the ensemble is in a pure state $e^i\omega|E_m\rangle$, the density operator $\rho = \Lambda_m$. Thus, from the above equation,

$$< A >_{\rho=\Lambda_m} = \mathrm{Tr}\,A\Lambda_m\ = \langle E_m|A|E_m\rangle\,,$$

which is the usual result for the expectation value of an operator in a pure state.

The mathematical quantity $< A >_\rho$ that is calculated from the theory corresponds to the average value $\langle A\rangle$ that is measured in experiments,

$$< A >_\rho = \mathrm{Tr}(A\rho)\ \dot{=} < A > \equiv \sum_n \frac{N_n}{N} a_n\ .$$

Similarly, the mathematical quantity $\Delta_\rho A$ calculated from the theory corresponds to the root-mean-square deviation Δa that is measured in experiments,

$$\Delta_\rho A = \sqrt{\mathrm{Tr}(\rho A^2) - (\mathrm{Tr}(\rho A))^2} \doteqdot \Delta a = \sqrt{\sum_i \frac{N_i a_i^2}{N} - (\sum_i \frac{N_i a_i}{N})^2}.$$

Statistics cannot in principle be eliminated from quantum mechanics: the outcome of the measurement of an observable is, in general, uncertain. For a quantum "particle", the uncertainty in its momentum ΔP and uncertainty in his position ΔQ satisfy

$$\Delta P \, \Delta Q \geq \frac{\hbar}{2} \, .$$

The above equation is known as the Heisenberg uncertainty relation and reveals that it is impossible to simultaneously know both the position and momentum with arbitrary precision.

A diatomic molecule comprised of two unlike atoms has an electric dipole moment. When the molecule oscillates, the charges that comprise the dipole moment are accelerated and, therefore, emit electromagnetic radiation. The average power radiated by a dipole as it makes a transition from the state $|E_m\rangle$ to $|E_n\rangle$ is

$$\overline{R}_{nm} = \frac{\omega_\gamma^4 \, q^2}{3\pi \epsilon_0 c^3} |\langle E_n | Q | E_m \rangle|^2 \, .$$

Because the dipole operator has non-zero matrix elements only between neighboring energy states, all transitions caused by the dipole operator are from the original energy state to one of the two adjacent states. This is called the selection rule for dipole transitions.

When a harmonic oscillator interacts with blackbody radiation, three transitions occur: spontaneous emission, induced emission and spontaneous absorption. For the transition $O_m \rightarrow O_n + \gamma$, the Einstein coefficient A_{mn} is the probability per unit time of spontaneous emission and is given by

$$A_{mn} = \frac{\omega_\gamma^3 \, q^2}{3\pi \epsilon_0 c^3 \hbar} |\langle E_n | Q | E_m \rangle|^2 \, .$$

The number of oscillators undergoing spontaneous emission per unit time is $N_m A_{mn}$, where N_m is the number of oscillators in the state $|E_m\rangle$. The probability per unit time of an oscillator undergoing induced emission is proportional to the density of radiation $u(\omega)$ at the angular frequency $\omega = (E_m - E_n)/\hbar$. The proportionality constant is defined to be the Einstein coefficient B_{mn}. Thus the number of oscillators undergoing induced emission per unit time is $N_m u(\omega) B_{mn}$. Finally, the probability per unit time of an oscillator undergoing spontaneous absorption is proportional to the density of radiation $u(\omega)$ at the angular frequency $\omega = (E_m - E_n)/\hbar$. The proportionality constant is defined to be the Einstein coefficient B_{nm}, so the number of oscillators undergoing spontaneous absorption per unit time is $N_n u(\omega) B_{nm}$. At equilibrium the number of oscillators undergoing emission per unit time must equal the number of oscillators undergoing absorption per unit time, which is satisfied if

$$B_{nm} = B_{mn} = \frac{\pi^2 c^3}{\hbar\omega^3} A_{mn} \,.$$

In the energy range where diatomic molecules can be described as harmonic oscillators, vibrating diatomic molecules emit or absorb electromagnetic radiation only with a single frequency. In addition, much weaker bands are observed that are caused by anharmonic forces. Finally, when the experimental resolution is increased, the single band that results from the harmonic force is found to consist of many equally-spaced bands.

Problems

For Sect. 1.2

1.1 Using the definitions for a and $a^\dagger$ in (1.2.16) and the Heisenberg commutation relations (1.2.11), derive (1.2.19) where $N = a^\dagger a$.

1.2 Using the definitions for a and $a^\dagger$ in (1.2.16) and the Heisenberg commutation relations (1.2.11), derive (1.2.20).

1.3 Using the definition $N = a^\dagger a$, and (1.2.20), verify (1.2.25a) and (1.2.25b).

1.4 Verify that if a and $a^\dagger$ satisfy $[a, a^\dagger] = 1$ and if N is defined by $N = a^\dagger a$, then a, $a^\dagger$, and N satisfy the following two equations:

$$[N, (a)^m] = -m(a)^m \;,\; [N, (a^\dagger)^m] = m(a^\dagger)^m \;, \quad m = \text{integer} \geq 0 \,.$$

1.5 Calculate the matrix element $\langle E_n | P | E_m \rangle$.

1.6

(a) Calculate the two matrix elements $\langle E_n | Q | E_n \rangle$ and $\langle E_n | P | E_n \rangle$.
(b) Calculate the two matrix elements $\langle E_n | Q^2 | E_n \rangle$ and $\langle E_n | P^2 | E_n \rangle$.
(c) Check your answers in (b) for consistency by using them to evaluate the matrix element of the Hamiltonian $H = \frac{1}{2\mu} P^2 + \frac{k}{2} Q^2$ between states $|E_n\rangle$.

1.7 Calculate the matrix element $\langle E_n | Q^2 | E_m \rangle$.

1.8 Find the energy levels of the one-dimensional Hamiltonian,

$$H = \frac{1}{2\mu}P^2 + \frac{1}{2}kQ^2 + k_1 1\,,$$

where k_1 is a constant.

1.9 Consider the one-dimensional Hamiltonian,

$$H = \frac{1}{2\mu}P^2 + \frac{1}{2}kQ^2 + k_1 Q + k_2 1 \quad ; \quad \mu, k, k_1, k_2 = \text{constants}\,.$$

(a) Show that H can be rewritten in the form

$$H = \frac{1}{2\mu}P^2 + \frac{1}{2}kQ'^2 + k_3 1\,,$$

where $Q' = Q + k_4\, 1$. Express the constants k_3 and k_4 in terms of k, k_1, and k_2.

(b) Calculate the commutator $[Q',\, P]$.

(c) Explain why the results from the analysis of the harmonic oscillator can now be used to solve the problem by making the substitution $Q \to Q'$.

(d) What are the energy levels E_n of H?

(e) Express the energy eigenstate $|E_n\rangle$ of H in terms of the operators P and Q' and the ground state $|E_0\rangle$.

1.10 Using the fact that the Hamiltonian is self-adjoint, show that $\langle E_n|E_m\rangle = 0$ for $n \neq m$. Hint: Obtain two expressions for the matrix element $\langle E_n|H|E_m\rangle$.

1.11 Using the expression for $|n\rangle = |E_n\rangle$ given in (1.2.48), derive (1.2.50).

For Sect. 1.3

1.12 Verify that the operator $\Lambda_n = |E_n\rangle\langle E_n|$ possesses the following properties that characterize a projection operator:

(a) $\Lambda_n^\dagger = \Lambda_n$

(b) $\Lambda_n \Lambda_n = \Lambda_n$

(c) $\Lambda_n \Lambda_m = 0$ for $n \neq m$

1.13 Let ψ and ϕ be normalized vectors satisfying $\langle\psi|\phi\rangle = 0$.

(a) Calculate $\Lambda_\psi \Lambda_\phi$ where $\Lambda_\psi \equiv |\psi\rangle\langle\psi|$ and $\Lambda_\phi \equiv |\phi\rangle\langle\phi|$.

(b) Calculate the probability of observing the state ψ in the mixture $\rho = \lambda\Lambda_\psi + (1-\lambda)\Lambda_\phi,\ 0 < \lambda < 1.$

1.14 Consider a mixture of quantum harmonic oscillators in states $|E_0\rangle$ and $|E_1\rangle$. Let λ and $1-\lambda$, $0<\lambda<1$, be the respective probabilities that an oscillator is in the state $|E_0\rangle$ or $|E_1\rangle$.

(a) Construct the statistical operator ρ for this mixture.
(b) Calculate $\langle H\rangle$ and $\langle Q\rangle$.
(c) Make the *incorrect* assumption that this mixture is described by the normalized vector,

$$\Psi = \sqrt{\lambda}\,|E_0\rangle + \sqrt{1-\lambda}\,|E_1\rangle\,,$$

and that the expectation value of an operator A in this mixture is given by $\langle\Psi|A|\Psi\rangle$. Calculate $\langle\Psi|H|\Psi\rangle$ and $\langle\Psi|Q|\Psi\rangle$.
(d) For what values of λ does the expectation value $\langle Q\rangle$ actually equal $\langle\psi|Q|\psi\rangle$? Explain why this result was to be expected.

1.15 For a linear operator A on the state ρ,

$$\rho = \sum_{n=0}^{\infty} w_n |E_n\rangle\langle E_n|\,,$$

show that $\mathrm{Tr}(A\rho) = \mathrm{Tr}(\rho A)$

For Sect. 1.4

1.16 The energy-loss spectrum of a vibrating H_2 molecules shown in Fig. 1.14 on the facing page has two peaks, one at 0 and the second at 0.52 eV with respective intensities of 3.5 and $7.8 \times 1/(30) = 0.26$ in arbitrary units. What is the statistical operator ρ for the ensemble of H_2 molecules?

1.17 Show that when $\rho = \Lambda_\phi = |\phi\rangle\langle\phi|$ is the statistical operator for a pure state, then the result for $\Delta_\phi A$ as given by (1.4.10) is identical to (1.4.9).

1.18 For the state $\rho = \lambda|E_n\rangle\langle E_n| + (1-\lambda)|E_m\rangle\langle E_m|,\ 0<\lambda<1,\ m\neq n$, calculate $\Delta_\rho H$ and verify that it is not zero.

1.19 For the state ρ given in (1.3.11) calculate

(a) $\Delta_\rho Q$
(b) and $\Delta_\rho P$.
(c) Restricting to the case $w_n = 1$, calculate $\Delta p \Delta x$.

1.20 Explain why the Heisenberg uncertainty principle would be violated if the zero-point energy of the harmonic oscillator were zero.

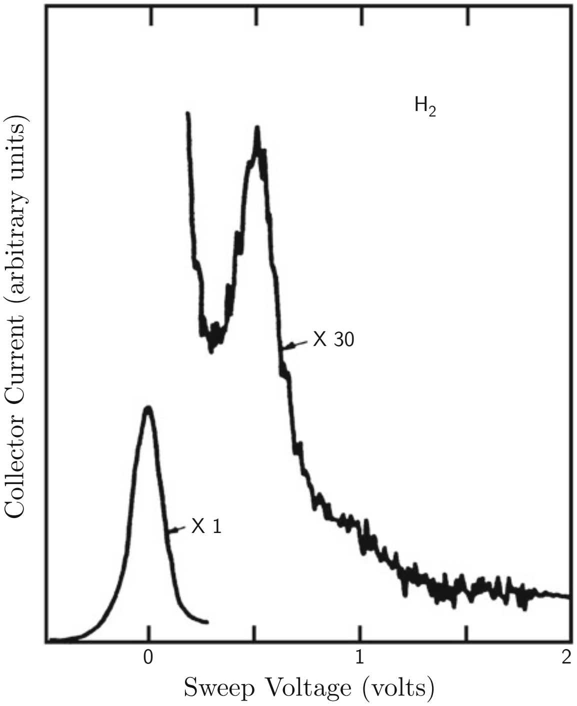


Fig. 1.14 Energy-loss spectrum of H_2. [From G. J. Schultz, Phys. Rev. **135**, A988,1964. with permission]

For Sect. 1.5

1.21 The vibrating HCl molecule emits infrared radiation with wavelength $\lambda = 3.46 \times 10^{-6}$ m.

(a) What is the ground state energy of the vibrating HCl molecule?
(b) What are the next three energy levels?

1.22 The transition probability for quadrupole transitions in a harmonic oscillator is proportional to the square of the magnitude of the matrix element $\langle E_n|Q^2|E_m\rangle$. What are the selection rules for quadrupole transitions?

1.23

(a) Show that $\overline{R}_{nm}$ has units of energy/time.
(b) Show that the Einstein coefficient A_{nm} has units of 1/time.
(c) Show that the Einstein coefficient B_{nm} multiplied by the energy density per angular frequency $u(\omega)$ has units of 1/time.

1.24 From the formula for B_{mn} in (1.5.30), show that $B_{mn} = B_{nm}$.

1.25 Using the explicit expressions for A_{mn} and B_{mn}, show that (1.5.25) is satisfied.

1.26 Answer the following questions about the CO molecule:

(a) What is its reduced mass μ?
(b) Using $E_{n+1} - E_n = 0.265\,\text{eV}$, calculate the "spring constant" k.
(c) What is the classical value of the square of the amplitude A^2 when the CO molecule is in the state $|E_n\rangle$?
(d) What is the expectation value of Q^2 in the state $|E_n\rangle$?

Chapter 2
Angular Momentum

2.1 Introduction

The diatomic molecule, which was considered in the previous chapter as an important example of a physical system that can be described as a harmonic oscillator, not only vibrates, but also rotates. The Hamiltonian H of a rotating diatomic molecule, which is the generator of time translations, commutes with all the position operators Q_i, the momentum operators P_i and the angular momentum operators J_i. These operators are generators of an algebra of observables. In this chapter attention is initially restricted to the subalgebra of the angular momentum operators J_i, the generators of the rotation group.

In Sect. 2.2 the quantum formula for orbital angular momentum is conjectured from the form of the classical formula. The Hamiltonian of a rotating, rigid dumbbell is similarly found. The representations of the algebra of angular momentum are determined in Sect. 2.3. In Sect. 2.4 the rotating diatomic molecule is discussed. While vibrations yield energy spectra with changes $\Delta E_{\text{vibration}} \approx 10^{-1}$ eV, for rotations the changes in energy are much smaller, $\Delta E_{\text{rotation}} \approx 10^{-3}$ eV. In Sect. 2.5 the calculation of probabilities in angular momentum states is considered, and in Sect. 2.6 the relationship between $SO(3)$ and $SU(2)$ is briefly discussed.

2.2 Introduction to Angular Momentum

In classical physics a mass point is a physical system characterized by one system parameter, the mass m, and two physical variables, the momentum **p** and the position **x**, both of which are measured with respect to a reference position usually denoted 0.

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9_2

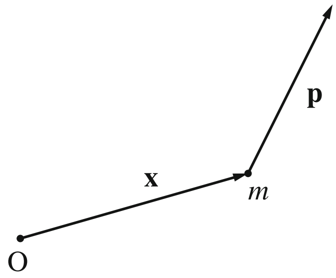


Fig. 2.1 The position vector **x** and momentum vector **p** for a mass-point *m*

The classical orbital angular momentum **l** is defined in terms of these two physical variables (Fig. 2.1),

$$\mathbf{l} = \mathbf{x} \times \mathbf{p}\,. \tag{2.2.1}$$

According to the correspondence principle, the quantum mechanical orbital angular momentum **L** of a quantum mechanical mass point is

$$\mathbf{L} = \mathbf{Q} \times \mathbf{P}, \tag{2.2.2}$$

where **Q** is the position operator, and **P** is the momentum operator. In component form,

$$L_x = Q_y P_z - Q_z P_y, \tag{2.2.3a}$$

$$L_y = Q_z P_x - Q_x P_z, \tag{2.2.3b}$$

$$L_z = Q_x P_y - Q_y P_x. \tag{2.2.3c}$$

Introducing the *permutation symbol*[1] ϵ_{ijk} defined by

$$\epsilon_{ijk} \equiv \begin{cases} 1 & \text{for every even permutation of } i,\ j \text{ and } k, \\ -1 & \text{for every odd permutation of } i,\ j \text{ and } k, \\ 0 & \text{if any two of the indices } i,\ j, \text{ and } k \text{ have the same value,} \end{cases} \tag{2.2.4}$$

all three equations (2.2.3) can be written compactly as the single equation,

$$L_i = \sum_{j,k=1}^{3} \epsilon_{ijk} Q_j P_k = \epsilon_{ijk} Q_j P_k\,. \tag{2.2.5}$$

---

[1] Also known as the Levi-Civita symbol or antisymmetric symbol.

For notational convenience summation symbols are often omitted with the understanding that any repeated index such as j and k in the above equation are to be summed over the range of the index. The final line in (2.2.5) adopts this notation.

From (2.2.4), $\epsilon_{123} = 1$. Then, for example, $\epsilon_{213} = -1$ because there has been a single permutation of the "2" and the "1". Also, for example, $\epsilon_{312} = 1$ because there have been two permutations: First the "3" and the "2" were permuted to yield ϵ_{132} and then the "3" and the "1" were permuted to yield ϵ_{312}. By using the permutation symbol, as many as nine equations can be written as a single equation, and calculations are often greatly simplified.

The commutation relations for the operators L_i can be calculated from the three-dimensional generalization of the canonical commutation relations for the momentum and position operator (1.2.11), namely,

$$[P_i, Q_j] = \frac{\hbar}{i}\delta_{ij}1, \quad [P_i, P_j] = 0, \quad [Q_i, Q_j] = 0, \quad i, j = x, y, z, \text{ or } 1, 2, 3. \tag{2.2.6}$$

Orbital angular momentum can be defined by $\mathbf{L} = \mathbf{Q} \times \mathbf{P}$ or , equivalently, by $\mathbf{L} = -\mathbf{P} \times \mathbf{Q}$. To see that the two expressions are identical, for the x-component of $\mathbf{L}$ note from (2.2.6) that $Q_y P_z = P_z Q_y$ and $Q_z P_y = P_y Q_z$. Then from (2.2.3a), the detailed calculation in terms of components is as follows:

$$\begin{aligned} L_x &= (\mathbf{Q} \times \mathbf{P})_x = Q_y P_z - Q_z P_y = (P_z Q_y - P_y Q_z) \\ &= -(P_y Q_z - P_z Q_y) = -(\mathbf{P} \times \mathbf{Q})_x \end{aligned} \tag{2.2.7}$$

Although the order in which the operators **P** and **Q** appear in the definition of orbital angular momentum is irrelevant, this is not the case when a quantum mechanical operator is defined in terms of two operators that do not commute.

As Q_i and P_j are hermitian, it follows that L_i is hermitian. For example using (2.2.3a),

$$L_x^\dagger = (Q_y P_z - Q_z P_y)^\dagger = P_z^\dagger Q_y^\dagger - P_y^\dagger Q_z^\dagger = P_z Q_y - P_y Q_z. \tag{2.2.8}$$

But, as was just mentioned, from the Heisenberg commutation relation (2.2.6) $P_z Q_y = Q_y P_z$ and $P_y Q_z = Q_z P_y$. Therefore,

$$L_x^\dagger = Q_y P_z - Q_z P_y = L_x. \tag{2.2.9}$$

In a similar fashion L_y and L_z can be shown to be hermitian.

The commutation relations between L_x, L_y and L_z follow from the definition of **L** and the Heisenberg commutation relations (2.2.6). For example,

$$\begin{aligned} [L_x, L_y] &= [Q_y P_z - Q_z P_y, Q_z P_x - Q_x P_z] \\ &= [Q_y P_z, Q_z P_x] - [Q_y P_z, Q_x P_z] - [Q_z P_y, Q_z P_x] + [Q_z P_y, Q_x P_z]. \end{aligned} \tag{2.2.10}$$

The second and third commutators are zero because all operators involved commute. In the first and fourth commutators only P_z and Q_z don't commute. Thus

$$[L_x, L_y] = Q_y P_x [P_z, Q_z] + Q_x P_y [Q_z, P_z]. \tag{2.2.11}$$

Using the Heisenberg commutation relation (2.2.6),

$$[L_x, L_y] = i\hbar (Q_x P_y - Q_y P_x) = i\hbar L_z, \tag{2.2.12a}$$

In a similar fashion the remaining two commutation relations are obtained,

$$[L_y, L_z] = i\hbar L_x, \tag{2.2.12b}$$

$$[L_z, L_x] = i\hbar L_y. \tag{2.2.12c}$$

Note that (2.2.12b) can be obtained from (2.2.12a) by making the substitution $x \to y, y \to z, z \to x$, and (2.2.12c) can be obtained from (2.2.12b) by making the same substitution. The substitution $x \to y, y \to z, z \to x$ is called "cyclically permuting the indices".

The three equations (2.2.12) can be written in the compact form

$$[L_i, L_j] = i\hbar\, \epsilon_{ijk} L_k\,, \tag{2.2.13}$$

where ϵ_{ijk} is defined in (2.2.4).

The commutation relations (2.2.13) for the orbital angular momentum of a mass point have been derived from the Heisenberg commutation relations (2.2.6). The orbital angular momentum $\mathbf{L} = \mathbf{Q} \times \mathbf{P}$ of the center of mass of a quantum mechanical rigid body will also fulfill (2.2.13) because, by the correspondence principle, the center of mass position $\mathbf{x} \to \mathbf{Q}$ and the center-of-mass momentum $\mathbf{p} = M\dot{x} \to \mathbf{P}$. That is, the classical center-of-mass position and center-of-mass momentum are replaced by the corresponding quantum operators. The operators $\mathbf{Q}$ and $\mathbf{P}$ fulfill the Heisenberg commutation relations because they represent, respectively, the position and momentum of the center of mass.

Classical *extended objects* can also possess a second type of angular momentum called spin: spin angular momentum $\mathbf{s}$ results from rotational motion about the center of mass while orbital angular momentum $\mathbf{l}$ results from motion of the center of mass about a fixed point. The total classical angular momentum $\mathbf{j}$ is then given by

$$\mathbf{j} = \mathbf{l} + \mathbf{s}. \tag{2.2.14}$$

From the correspondence principle, the above equation is also valid for quantum mechanical systems provided the classical quantities are replaced by the corresponding quantum operators,

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \tag{2.2.15}$$

where **J** and **S** are, respectively, the quantum mechanical operators for total angular momentum and spin angular momentum.[2]

2.2.1 *Rigid Dumbbell as a Model of a Rotating Diatomic Molecule*

A simple example of a physical system with rotational degrees of freedom is the rigid dumbbell. The classical picture of a dumbbell is two mass points $m^{(1)}$ and $m^{(2)}$ at opposite ends of a massless rod. A dumbbell represents a diatomic molecule, which is a quantum system. The two mass points are the nuclei of the two atoms that make up the diatomic molecule, and the massless rod along the internuclear axis represents the atomic (electromagnetic) forces that keep the two nuclei at a fixed distance from each other.

If sufficient energy is supplied to a diatomic molecule, the molecule can *vibrate* as well as *rotate*, implying that it is no longer a rigid rotator. Then the rigid rod instead acts like a rotating spring connecting the two masses, and the molecule both vibrates and rotates, thus becoming a vibrating rotator. If even more energy is supplied, single electrons will perform transitions between electronic energy levels. As increasing energy is supplied, new degrees of freedom are manifest.

The rigid rotator model of the diatomic molecule is valid for energies in the infrared region that are *incapable* of inducing transitions to excited vibrational or excited electronic states. It is an empirical fact that only rotational degrees of freedom can be excited at energies below about 10^{-2} eV so attention is initially restricted to the rigid dumbbell.

The algebra of the quantum mechanical observables for the rigid dumbbell is conjectured from correspondence to the observables for the classical, rigid rotator. First the relationship between energy and angular momentum is derived. Then the classical-quantum correspondence is used to obtain the quantum Hamiltonian.

Since only rotations about the center of mass are considered, the center of mass is assumed to be fixed at the origin, and the positions of the point masses $m^{(1)}$ and $m^{(2)}$ are measured from the center of mass. Thus it is convenient to measure the motion of the molecule in its center-of-mass frame. The position vectors $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ for masses m_1 and m_2, respectively, are shown in Fig. 2.2 on the next page.

The total energy E of the two particles is the sum of the kinetic energies of each,

$$E = \frac{1}{2}m^{(1)}\left[\frac{\mathrm{d}\mathbf{x}^{(1)}}{\mathrm{d}t}\right]^2 + \frac{1}{2}m^{(2)}\left[\frac{\mathrm{d}\mathbf{x}^{(2)}}{\mathrm{d}t}\right]^2 . \tag{2.2.16}$$

---

[2] In quantum mechanics spin is an intrinsic property of a particle, unrelated to any sort of motion in space.

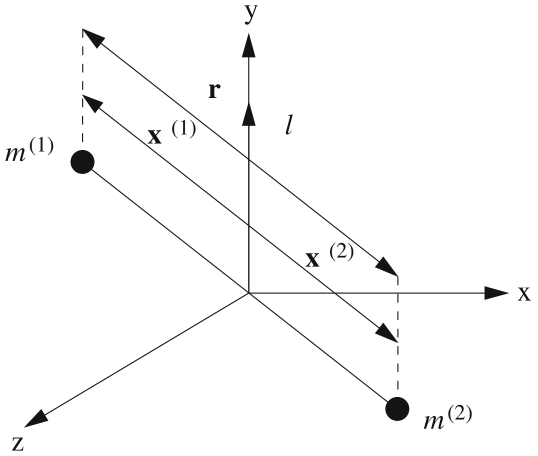


Fig. 2.2 The position vectors $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ for masses $m^{(1)}$ and $m^{(2)}$, respectively, and the vector $\mathbf{r} = \mathbf{x}^{(1)} - \mathbf{x}^{(2)}$, which is the position of $m^{(1)}$ with respect to $m^{(2)}$, for a rigid diatomic molecule

The angular momentum $\mathbf{l}$ of the dumbbell about the center of mass is the sum of the orbital angular momenta of the two mass points:

$$\mathbf{l} = \mathbf{l}^{(1)} + \mathbf{l}^{(2)} = \mathbf{x}^{(1)} \times m^{(1)} \frac{\mathrm{d}\mathbf{x}^{(1)}}{\mathrm{d}t} + \mathbf{x}^{(2)} \times m^{(2)} \frac{\mathrm{d}\mathbf{x}^{(2)}}{\mathrm{d}t} . \tag{2.2.17}$$

It is convenient to express the energy E and angular momentum $\mathbf{l}$ in terms of the vector

$$\mathbf{r} = \mathbf{x}^{(1)} - \mathbf{x}^{(2)}, \tag{2.2.18}$$

depicted in Fig. 2.2. This is accomplished by first expressing $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ in terms of $\mathbf{r}$ and $\mathbf{x}_{\text{CM}} = 0$.

Example 2.2.1 Find the center of mass $\mathbf{x}_{\text{CM}}$ of the diatomic molecule depicted in Fig. 2.2, but do not assume that the center of mass is at the origin. Express the vectors $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ in terms of $m^{(1)}$, $m^{(2)}$, $\mathbf{r}$ and $\mathbf{x}_{\text{CM}}$.

Solution The center of mass $\mathbf{x}_{\text{CM}}$ is calculated from the equation

$$\mathbf{x}_{\text{CM}} = \frac{\sum_{i=1}^{N} m^{(i)}\mathbf{x}^{(i)}}{\sum_{i=1}^{N} m^{(i)}} = \frac{m^{(1)}\mathbf{x}^{(1)} + m^{(2)}\mathbf{x}^{(2)}}{m^{(1)} + m^{(2)}}. \tag{2.2.19}$$

Solving the simultaneous linear equations (2.2.18) and (2.2.19) for $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ yields,

$$\mathbf{x}^{(1)} = \mathbf{x}_{\text{CM}} + \frac{m^{(2)}}{m^{(1)} + m^{(2)}}\,\mathbf{r} , \qquad \mathbf{x}^{(2)} = \mathbf{x}_{\text{CM}} - \frac{m^{(1)}}{m^{(1)} + m^{(2)}}\,\mathbf{r} . \tag{2.2.20}$$

Choosing the coordinate system depicted in Fig. 2.2 on the facing page with the center of mass at the origin, implying that $\mathbf{x}_{\mathrm{CM}} = 0$, and using the expressions (2.2.20), the formula for energy (2.2.16) can be rewritten in terms of $\mathbf{r}$,

$$E = \frac{1}{2}m^{(1)}\left[\frac{m^{(2)}}{m^{(1)}+m^{(2)}}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right]^2 + \frac{1}{2}m^{(2)}\left[-\frac{m^{(1)}}{m^{(1)}+m^{(2)}}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right]^2 = \frac{1}{2}\mu\left[\frac{\mathrm{d}r}{\mathrm{d}t}\right]^2 , \tag{2.2.21}$$

where μ is the reduced mass

$$\mu = \frac{m^{(1)}m^{(2)}}{m^{(1)}+m^{(2)}}. \tag{2.2.22}$$

In a similar calculation the orbital angular momentum $\mathbf{l}$ as given in (2.2.17) can also be expressed in terms of $\mathbf{r}$,

$$\begin{aligned}\mathbf{l} = {} & \frac{m^{(2)}}{m^{(1)}+m^{(2)}}\mathbf{r} \times m^{(1)}\left[\frac{m^{(2)}}{m^{(1)}+m^{(2)}}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right] \\ & - \frac{m^{(1)}}{m^{(1)}+m^{(2)}}\mathbf{r} \times m^{(2)}\left[-\frac{m^{(1)}}{m^{(1)}+m^{(2)}}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right] = \mu\mathbf{r} \times \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} .\end{aligned} \tag{2.2.23a}$$

or

$$l_i = \mu\epsilon_{ijk}r_j\frac{\mathrm{d}\mathbf{r}_k}{\mathrm{d}t} . \tag{2.2.23b}$$

With the aid of the identity[3]

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km} , \tag{2.2.24}$$

and the fact that the motion is purely rotational so that $\mathbf{r} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = 0$,

$$l_i l_i = \mathbf{l}^2 = \mu^2\left[(\mathbf{r}\cdot\mathbf{r})(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}) - (\mathbf{r}\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t})^2\right] = \mu^2 r^2\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 . \tag{2.2.25}$$

The moment of inertia I about the center of mass is defined by

$$I = m^{(1)}[x^{(1)}]^2 + m^{(2)}[x^{(2)}]^2. \tag{2.2.26}$$

---

[3] Multiplying (2.2.24) by $A_jB_kA_mB_n$ yields the vector identity

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{A}\times\mathbf{B}) = (\mathbf{A}\cdot\mathbf{A})(\mathbf{B}\cdot\mathbf{B}) - (\mathbf{A}\cdot\mathbf{B})^2.$$

If, for example, the diatomic molecule is HCl, then $m^{(1)} = m_{\mathrm{H}}$ and $m^{(2)} = m_{\mathrm{Cl}}$. Using (2.2.20) and recalling that $\mathbf{x}_{\mathrm{CM}} = 0$,

$$I = m^{(1)} \left(\frac{m^{(2)}}{m^{(1)} + m^{(2)}} \mathbf{r} \right)^2 + m^{(2)} \left(-\frac{m^{(1)}}{m^{(1)} + m^{(2)}} \mathbf{r} \right)^2 = \mu r^2. \tag{2.2.27}$$

Equations (2.2.25) and (2.2.27) allow the expression (2.2.21) for energy to be written in the desired form,

$$E = \frac{\mathbf{l}^2}{2I} . \tag{2.2.28}$$

To obtain the mathematical theory for the rigid, quantum dumbbell, the fundamental classical observables are replaced by the corresponding quantum operators:

Classical		Quantum	
Energy E	$\longrightarrow$	Hamiltonian operator H ,	(2.2.29a)
dumbbell axis $\mathbf{r}$	$\longrightarrow$	position operator $\mathbf{Q}$,	(2.2.29b)
angular momentum $\mathbf{l}$	$\longrightarrow$	angular momentum operator $\mathbf{L}$.	(2.2.29c)

The relationships among the quantum operators correspond to the classical relationships. Replacing the classical quantities with the corresponding quantum operators as indicated in (2.2.29), it follows from (2.2.28) that the Hamiltonian of the quantum rotator is

$$E = \frac{\mathbf{l}^2}{2I} \longrightarrow H = \frac{\mathbf{L}^2}{2I} . \tag{2.2.30}$$

The classical observables $\mathbf{l}$ and $\mathbf{r}$ are replaced by the quantum operators $\mathbf{L}$ and $\mathbf{Q}$, respectively, that fulfill the relations,

$$[Q_i, Q_j] = 0 , \qquad [L_i, L_j] = i\hbar\epsilon_{ijk} L_k , \qquad [L_i, Q_j] = i\hbar\epsilon_{ijk} Q_k . \tag{2.2.31}$$

The first of the above relations was postulated in (2.2.6), the second equation is (2.2.13), and the third equation is readily derived from (2.2.5) and (2.2.6). (See Problem 2.2.)

If the dumbbell molecule is rigid, implying that the molecule does not vibrate, the length r of the internuclear axis (2.2.18) must be constant:

$$\mathbf{Q}^2 \equiv Q_i Q_i = r^2 = \text{constant}. \tag{2.2.32}$$

Using (2.2.5), it follows that

$$\mathbf{Q} \cdot \mathbf{L} = Q_i L_i = \epsilon_{ijk} Q_i Q_j P_k \,. \tag{2.2.33}$$

Using $\epsilon_{ijk} = -\epsilon_{jik}$ and (2.2.6), which implies that $Q_i Q_j = Q_j Q_i$,

$$\mathbf{Q} \cdot \mathbf{L} = -\epsilon_{jik} Q_j Q_i P_k \,. \tag{2.2.34}$$

Making the change $i \to j$, $j \to i$ in the summation indices, the above equation becomes $\mathbf{Q} \cdot \mathbf{L} = -\epsilon_{ijk} Q_i Q_j P_k = -\mathbf{Q} \cdot \mathbf{L}$. Therefore,

$$\mathbf{Q} \cdot \mathbf{L} = 0 \,. \tag{2.2.35}$$

Equation (2.2.35) reveals that for the quantum dumbbell consisting of two mass points that do not have any intrinsic angular momentum (spin), the operator $\mathbf{Q}$ for the internuclear distance is orthogonal to the angular momentum operator $\mathbf{L}$. Therefore, the quantum mechanical observables Q_i and L_i describing a rigid, quantum dumbbell consisting of two spinless mass points fulfill the two relations,

$$\mathbf{Q}^2 = r^2 = \text{constant}, \quad \mathbf{Q} \cdot \mathbf{L} = 0 \,. \tag{2.2.36}$$

In Fig. 2.2 on page 70, which depicts the analogous classical system, the internuclear distance $\mathbf{r}$ is orthogonal to the orbital angular momentum $\mathbf{l}$.

2.3 Representations of the Algebra of Angular Momentum $\mathscr{E}$(SO(3))

In (2.2.2) the orbital angular momentum operators $\mathbf{L}$ were defined in terms of the position operator $\mathbf{Q}$ and the momentum operator $\mathbf{P}$. With the aid of the Heisenberg commutation relations (2.2.6), it was then possible to calculate the commutation relations (2.2.13) for the L_i. However, for an extended object, in addition to the orbital angular momentum L_i about a fixed point, there are other angular momentum operators. The fundamental conserved quantity is the total angular momentum J_i. The J_i cannot be defined in terms of P_i and Q_i, but since they are angular momentum operators, they are *defined* as the linear, self-adjoint operators that obey the relations

$$[J_i, J_j] = i\hbar \, \epsilon_{ijk} J_k \,. \tag{2.3.1}$$

The above relation identifies the angular momentum operators J_i as the generators of the group of rotations in three-dimensional space. The operators P_i, Q_i, and the Hamiltonian H are also generators of a group of motions, the Galilei group that is the group of motions in non-relativistic, three-dimensional space-time. The angular momentum operators are the generators of three-dimensional rotations in a subgroup of the Galilei group. The commutation relations (2.3.1) are a consequence of the property of the three-dimensional space as are the commutation relations of all the operators P_i, Q_i, H and J_i.

General angular momentum operators are defined to be the linear, self-adjoint operators $J_i = J_i^\dagger$ that fulfill (2.3.1). Their properties are derived from the commutation relations (2.3.1). Since quantum mechanical observables are operators in a linear, scalar-product space $\mathscr{H}$ or Φ, the objective of the derivation is to find all possible spaces and the properties of the linear operators J_i in these spaces.

The operator

$$J_i J_i \equiv \mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2 \,, \tag{2.3.2}$$

is the square of the total angular momentum and has the property (See Problem 2.3.)

$$[\mathbf{J}^2, J_i] = 0 \,. \tag{2.3.3}$$

Allowed values of angular momentum are determined for those J_i acting in spaces that have at least one eigenvector $|c\rangle$ of the operator $\mathbf{J}^2$ for which

$$\mathbf{J}^2|c\rangle = \hbar^2 c|c\rangle \,. \tag{2.3.4}$$

It is possible to choose a vector $|c\rangle$ in these spaces such that it is an eigenvector of both $\mathbf{J}^2$ and a second operator $\tilde{J}$, which is a linear combination of the three J_i. This eigenvector of both $\mathbf{J}^2$ and $\tilde{J}$ is denoted $|c, m\rangle$ and satisfies

$$\mathbf{J}^2|c, m\rangle = \hbar^2 c|c, m\rangle \,, \tag{2.3.5a}$$

and

$$\tilde{J}|c, m\rangle = \hbar m|c, m\rangle \,. \tag{2.3.5b}$$

As a result of the factor of $\hbar$ in (2.3.1), which occurred because of the factor of $\hbar$ in (2.2.6) that lead to the factors of $\hbar$ in (2.2.31), angular momentum is measured in units of $\hbar$. Powers of $\hbar$ are included in (2.3.5a) and (2.3.5b) so that the parameters c and m are dimensionless.

The quantity $\hbar^2 c$ is called the eigenvalue of $\mathbf{J}^2$ and $\hbar m$ is called the eigenvalue of $\tilde{J}$. If an operator $\tilde{J}$ exists, then, in general, the eigenvalue c will not uniquely specify the vector because the eigenvalue m is also required. Thus the desire to uniquely specify eigenvectors forces a search for an operator $\tilde{J}$. To determine the relationship between $\mathbf{J}^2$ and $\tilde{J}$, the commutator $[\mathbf{J}^2, \tilde{J}]$ is applied to $|c, m\rangle$, and then (2.3.5a) and (2.3.5b) are used:

$$\begin{aligned}[\mathbf{J}^2, \tilde{J}]|c, m\rangle &= \mathbf{J}^2\tilde{J}|c, m\rangle - \tilde{J}\mathbf{J}^2|c, m\rangle \\ &= \mathbf{J}^2\hbar m|c, m\rangle - \tilde{J}\hbar^2 c|c, m\rangle = \hbar^3(cm - cm)|c, m\rangle = 0 \,.\end{aligned} \tag{2.3.6}$$

Therefore, if $|c, m\rangle$ is an eigenvector of both $\mathbf{J}^2$ and $\tilde{J}$, the two operators must commute when acting on any vector $|c, m\rangle$.

Because $\mathbf{J}^2$ commutes with J_i, $\tilde{J}$ can be chosen to be the component of $\mathbf{J}$ along an arbitrary unit vector $\hat{\mathbf{n}}$ in three-dimensional space, $\tilde{J} = \hat{\mathbf{n}} \cdot \mathbf{J}$. It is customary to call the direction of $\hat{\mathbf{n}}$ the three-direction so that $\hat{\mathbf{n}} = \mathbf{e}_3$ and $\tilde{J} = \hat{\mathbf{n}} \cdot \mathbf{J} = \mathbf{e}_3 \cdot \mathbf{J} = J_3$. As a consequence, $|c, m\rangle$ is chosen to be an eigenstate of the component of $\mathbf{J}$ called J_3,

$$\mathbf{J}^2|c, m\rangle = \hbar^2 c|c, m\rangle\,, \tag{2.3.7a}$$

$$J_3|c, m\rangle = \hbar m|c, m\rangle\,. \tag{2.3.7b}$$

When rotational symmetry is destroyed by introducing an external electric or magnetic field, for example, a specific direction in space has significance. Under these circumstances it is often convenient to choose e_3 in the direction of the external field. Otherwise the direction of e_3 is arbitrary.

If a third operator existed that was linearly independent of $\mathbf{J}^2$ and J_3 and also commuted with both operators, then the eigenvalue of this third operator would also be required to uniquely specify the eigenvector. However, no such operator can be constructed from just the J_i so the eigenvectors are completely specified by the eigenvalues c and m. Because the operators $\mathbf{J}^2$ and J_3 are self-adjoint, the parameters c and m are real.

The set of operators

$$A = a_0 + \sum_{i=1}^{3} a_i J_i + \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij} J_i J_j + \ldots \qquad a_i, a_{ij}, \dots \in \mathbb{C}\,, \tag{2.3.8}$$

forms an algebra of linear operators as discussed in Sect. A.3 on page 320 of the Appendix. The algebra of all operators of the form (2.3.8) generated by the J_i is called the enveloping algebra of the rotation group SO(3) and is denoted $\mathscr{E}(\mathrm{SO}(3))$.

Since $\mathbf{J}^2$ commutes with all the operators J_i, it follows that $\mathbf{J}^2$ also commutes with all A:

$$[A, \mathbf{J}^2] = 0 \ \text{ for all } A \in \mathscr{E}(\mathrm{SO}(3)) \tag{2.3.9}$$

From (2.3.9) it follows that there exists no $A \in \mathscr{E}(\mathrm{SO}(3))$ that changes the value of c. This can be shown by applying $[\mathbf{J}^2, A] = 0$ to the vector $|c, m\rangle$:

$$\mathbf{J}^2 A|c, m\rangle = A\mathbf{J}^2|c, m\rangle = \hbar^2 c A|c, m\rangle \tag{2.3.10}$$

From the first and last term in (2.3.10), it follows that the vector $A|c, m\rangle$ is again an eigenvector of $\mathbf{J}^2$ with the same eigenvalue $\hbar^2 c$. This means that the set of all vectors $\{\psi^c = A|c, m\rangle\}$ where $A \in \mathscr{E}(\mathrm{SO}(3))$ (i.e. where A is any operator of the form (2.3.8) with arbitrary coefficients $a_0, a_j, a_{ij}, \dots \in \mathbb{C}$) is again a set of eigenvectors of $\mathbf{J}^2$ with the same eigenvalue c, $\mathbf{J}^2\psi^c = c\psi^c$.

There are operators that don't commute with J_3 and thus change the eigenvalue $\hbar m$ of J_3. An operator Ω is now sought that transforms an eigenvector $|c, m\rangle$ of J_3 with eigenvalue $\hbar m$ into an eigenvector of J_3 with eigenvalue $\hbar(m + \lambda)$. Such an operator must satisfy the relation

$$J_3 \Omega |c, m\rangle = \hbar(m + \lambda)\Omega |c, m\rangle \,, \tag{2.3.11}$$

and will make it possible to construct the eigenvectors $|c, m + \lambda\rangle$, $|c, m + 2\lambda\rangle$ etc. from the eigenvector $|c, m\rangle$. Multiplying (2.3.7b) by Ω,

$$\Omega J_3 |c, m\rangle = \hbar m \Omega |c, m\rangle \,, \tag{2.3.12}$$

and then subtracting the result from (2.3.11) yields

$$[J_3, \Omega]|c, m\rangle = \hbar\lambda\Omega |c, m\rangle \,. \tag{2.3.13}$$

The above equation is satisfied if J_3 and Ω satisfy the operator relation

$$[J_3, \Omega] = \hbar\lambda\Omega \,. \tag{2.3.14}$$

If the operator Ω is a linear function of the J_i, it must be of the form $\Omega = J_1 + \alpha J_2 + \beta J_3$ where α and β are constants and the coefficient of J_1 has been set equal to unity for convenience. Using the commutation relations (2.3.1), Eq. (2.3.14) implies

$$i\hbar J_2 - i\hbar\alpha J_1 = \hbar\lambda(J_1 + \alpha J_2 + \beta J_3) \,, \tag{2.3.15}$$

or

$$0 = \hbar(\lambda + i\alpha)J_1 + \hbar(\lambda\alpha - i)J_2 + \hbar\lambda\beta J_3 \,. \tag{2.3.16}$$

Since the J_i are linearly independent, it is not possible to express one operator in terms of the other two, implying that the above equation can only be satisfied if the coefficient of each J_i vanishes independently. Thus

$$0 = \lambda + i\alpha \,, \quad 0 = \lambda\alpha - i \,, \quad \beta = 0 \,. \tag{2.3.17}$$

The first equation in (2.3.17) implies $\lambda = -i\alpha$. Substituting this expression for λ into the second equation in (2.3.17) yields the two solutions to (2.3.17):

$$\alpha = i \,, \qquad \lambda = 1 \,, \tag{2.3.18a}$$

$$\alpha = -i \,, \qquad \lambda = -1 \,. \tag{2.3.18b}$$

The operators corresponding to the above two solutions are, respectively,

$$\Omega|_{\alpha=i} = J_1 + iJ_2 \equiv J_+ \,, \tag{2.3.19a}$$

$$\Omega|_{\alpha=-i} = J_1 - iJ_2 \equiv J_- \,. \tag{2.3.19b}$$

Because J_1 and J_2 are hermitian, it follows that the operators $J_\pm$ satisfy

$$J_\pm^\dagger = J_1^\dagger \mp iJ_2^\dagger = J_1 \mp iJ_2 = J_\mp \,. \tag{2.3.20}$$

Alternatively, *spherical components* J_0 and $J_{\pm 1}$ of the angular momentum vector operator are sometimes used. These operators are defined in terms of the J_i that fulfill (2.3.1):

$$J_0 \equiv J_3 \,, \tag{2.3.21a}$$

$$J_{\pm 1} \equiv \mp \frac{1}{\sqrt{2}}(J_1 \pm iJ_2) \,. \tag{2.3.21b}$$

The spherical components of angular momentum fulfill the relations

$$J_{\pm 1}^\dagger = -J_{\mp 1}, \qquad J_0^\dagger = J_0 \,. \tag{2.3.22}$$

From (2.3.11) and (2.3.19a), when $\lambda = 1$ in (2.3.18a), the operator $J_+ = J_1 + iJ_2$ increases the three-component of the angular momentum by one unit of angular momentum $\hbar$ and is therefore called a raising operator. Similarly $J_- = J_1 - iJ_2$ decreases the three-component of the angular momentum by $\hbar$ and is called a lowering operator.

From the algebraic solution of the angular momentum, it was possible to guess that the operators that change the three-component of angular momentum are non-hermitian. In fact J_+ and J_- are analogous to the operators $a^\dagger$ and a of the harmonic oscillator in that J_+ and $a^\dagger$ are both raising operators and J_- and a are both lowering operators. From (2.3.1) and (2.3.19) it immediately follows that

$$[J_3, J_\pm] = \pm\hbar J_\pm \,. \tag{2.3.23}$$

The allowed values of the eigenvalues c and m can now be determined. Because the eigenvalue c is fixed, a constraint is placed on the values that m can take. Using the Hermiticity of J_i, $J_i = J_i^\dagger$,

$$\begin{aligned}\langle c, m|J_1^2 + J_2^2|c, m\rangle &= \langle c, m|J_1^2|c, m\rangle + \langle c, m|J_2^2|c, m\rangle \\ &= (J_1|c, m\rangle, J_1|c, m\rangle) + (J_2|c, m\rangle, J_2|c, m\rangle) \geq 0 \,. \end{aligned} \tag{2.3.24}$$

The inequality follows because the scalar product of any vector $|d\rangle$ satisfies $(|d\rangle, |d\rangle) \equiv \langle d|d\rangle \geq 0$. Thus from (2.3.24) it follows that the expectation value (or eigenvalue) of the operator $J_1^2 + J_2^2$ is equal to or greater than zero. Allowing $J_1^2 + J_2^2$ to act on the vector $|c, m\rangle$,

$$\begin{aligned}(J_1^2 + J_2^2)|c, m\rangle = (J_1^2 + J_2^2 + J_3^2 - J_3^2)|c, m\rangle &= (\mathbf{J}^2 - J_3^2)|c, m\rangle \\ &= \hbar^2(c - m^2)|c, m\rangle \,. \end{aligned} \qquad (2.3.25)$$

Comparing (2.3.24) and (2.3.25) yields the constraint $(c - m^2) \geq 0$. To avoid violating this inequality, there must be a maximum and minimum value of m, denoted, respectively, by m_{max} and m_{min}. Note that this conclusion followed from the Hermiticity of J_i. This means that there exists no eigenvector of J_3 with an eigenvalue larger than m_{max} or smaller than m_{min}, i.e.,

$$J_+|c, m_{\text{max}}\rangle = 0\,, \qquad (2.3.26a)$$

$$J_-|c, m_{\text{min}}\rangle = 0\,. \qquad (2.3.26b)$$

Using the easily derived relation (Problem 2.4),

$$J_\pm J_\mp = \mathbf{J}^2 - J_3^2 \pm \hbar J_3\,, \qquad (2.3.27)$$

and allowing J_- and J_+ to act on (2.3.26a) and (2.3.26b), respectively,

$$0 = J_- J_+|c, m_{\text{max}}\rangle = \hbar^2(c - m_{\text{max}}^2 - m_{\text{max}})|c, m_{\text{max}}\rangle\,, \qquad (2.3.28a)$$

$$0 = J_+ J_-|c, m_{\text{min}}\rangle = \hbar^2(c - m_{\text{min}}^2 + m_{\text{min}})|c, m_{\text{min}}\rangle\,. \qquad (2.3.28b)$$

Equations (2.3.28a) and (2.3.28b) imply

$$c = m_{\text{max}}^2 + m_{\text{max}} = m_{\text{min}}^2 - m_{\text{min}}\,, \qquad (2.3.29)$$

or

$$(m_{\text{max}} + m_{\text{min}})(m_{\text{min}} - m_{\text{max}} - 1) = 0\,. \qquad (2.3.30)$$

Now $m_{\text{min}} - m_{\text{max}} \leq 0$, so the second term in (2.3.30) is unequal to zero. As a consequence the only solution is

$$m_{\text{max}} = -m_{\text{min}} \equiv j\,, \qquad (2.3.31)$$

where m_{max} is now denoted by j. From (2.3.29) it then follows that

$$c = j(j+1)\,. \qquad (2.3.32)$$

Instead of labeling eigenvectors by c, they will now be labeled by j, which is called the angular momentum quantum number.

What are the allowed values for j? The parameter $m_{\text{max}} = j$, and the lowering operator $\hat{J}_-$ decreases m in integer increments until it has a value $m_{\text{min}} = -j$. Since the lowering operator must be applied n times to transform $|j, m_{\text{max}}\rangle$ into $|j, m_{\text{min}}\rangle$ then

$$m_{\text{max}} - n = m_{\text{min}} \,. \tag{2.3.33}$$

Combining the above equation with (2.3.31), it follows that $j = n/2$. That is, j can take any integer or half-integer values. For a specific value of j, the allowed values of m are

$$m = j, j-1, j-2, \ldots, -j+1, -j \,. \tag{2.3.34}$$

Thus for every integer or half-integer value j, there exists a set of $(2j+1)$ eigenvectors with the property

$$\mathbf{J}^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle, \quad J_3|j,m\rangle = \hbar m|j,m\rangle \,. \tag{2.3.35}$$

The J_i act in the $(2j+1)$-dimensional space spanned by the ψ^j,

$$\{\psi^j = \sum_{m=-j}^{m=j} a_m|j,m\rangle, \quad a_m \in \mathbb{C}\} \,, \tag{2.3.36}$$

where the $a_m = \langle j,m|\psi^j\rangle$ are the expansion coefficients or components of the vector ψ^j.

In Table 2.1 allowed values of m are given for various values of j. When $j = 2$, for example, the allowed values of m are displayed in Fig. 2.3 on the following page. Each dot in the figure represents a one-dimensional space spanned by the vector $|j = 2, m\rangle$.

Eigenvectors of a hermitian operator with different eigenvalues are orthogonal. Since $\mathbf{J}^2$ and J_3 are hermitian,

$$\langle j,m|j,m'\rangle = \delta_{m,m'} \,, \tag{2.3.37}$$

Table 2.1 Allowed values of the third component of angular momentum m for j = 1/2, 1, 3/2 and 2

j	m	j	m	j	m	j	m
1/2	1/2	1	1	3/2	3/2	2	2
	−1/2		0		1/2		1
			−1		−1/2		0
					−3/2		−1
							−2

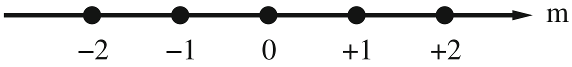


Fig. 2.3 Weight diagram with $j = 2$ for an irreducible representation of SO(3)

where the eigenvectors have been chosen to have unit length.

Matrix elements of the operators $J_{\mp}$ in (2.3.19) will now be determined. From (2.3.11) it follows that when $\Omega = J_{\mp}$ acts on the basis vector $|j, m\rangle$, it creates a vector that is an eigenvector of J_3 with an eigenvalue $\hbar(m \pm 1)$. Thus the vector is proportional to $|j, m \mp 1\rangle$:

$$J_{\mp}|j, m\rangle = \hbar\alpha_{\mp}|j, m \mp 1\rangle \,, \tag{2.3.38}$$

where the proportionality constants $\alpha_{\mp}$ depend on j and m and must be chosen so that the eigenvectors $|j, m \mp 1\rangle$ are normalized according to (2.3.37). From (2.3.38),

$$\begin{aligned}\hbar^2|\alpha_{\mp}|^2 &= (\hbar\alpha_{\mp}|j, m \mp 1\rangle, \hbar\alpha_{\mp}|j, m \mp 1\rangle = (J_{\mp}|j, m\rangle, J_{\mp}|j, m\rangle) \\ &= \langle j, m|J_{\mp}^{\dagger} J_{\mp}|j, m\rangle.\end{aligned} \tag{2.3.39}$$

Equations (2.3.20) and (2.3.27) then allow the calculation of $\alpha_{\mp}$ from (2.3.39):

$$\begin{aligned}\hbar^2|\alpha_{\mp}|^2 &= \langle j, m|\mathbf{J}^2 - J_3^2 \pm \hbar J_3|j, m\rangle = \hbar^2[j(j+1) - m^2 \pm m] \\ &= \hbar^2(j \pm m)(j \mp m + 1) \,.\end{aligned} \tag{2.3.40}$$

Except for a phase, which is chosen to be unity,

$$\alpha_{\mp} = \sqrt{(j \pm m)(j \mp m + 1)} \,. \tag{2.3.41}$$

Example 2.3.1 Express $J_1|j, m\rangle$ in terms of normalized eigenvectors.

Solution Using (2.3.19) to write J_1 in terms of J_+ and J_- and then using (2.3.38) and (2.3.41),

$$J_1|j, m\rangle = \frac{1}{2}(J_+ + J_-)|j, m\rangle = \frac{1}{2}\hbar\alpha_+|j, m+1\rangle + \frac{1}{2}\hbar\alpha_-|j, m-1\rangle \,,$$

or

$$J_1|j, m\rangle = \frac{\hbar}{2}\sqrt{(j-m)(j+m+1)}|j, m+1\rangle + \frac{\hbar}{2}\sqrt{(j+m)(j-m+1)}|j, m-1\rangle \,.$$

Example 2.3.2 Calculate the matrix element $\langle j, m|J_1^2|j, m\rangle$.

Solution Again using (2.3.19) to express J_1 in terms of J_+ and J_-,

$$\langle j, m|J_1^2|j, m\rangle = \langle j, m|\frac{1}{2}(J_+ + J_-)\frac{1}{2}(J_+ + J_-)|j, m\rangle\,.$$

The only non-zero contributions will come from operator products that contain both a raising and a lowering operator because both original eigenvectors have the same value of m.

$$\langle j, m|J_1^2|j, m\rangle = \langle j, m|\frac{1}{4}(J_+J_- + J_-J_+)|j, m\rangle = \frac{\hbar^2}{2}[j(j+1) - m^2]\,.$$

The last line can be calculated most easily from (2.3.27) although (2.3.38) and (2.3.41) could be used instead.

For every integer and half-integer value $j = 0,\ 1/2,\ 1,\ 3/2,\ \ldots$, there exists a $(2j+1)$-dimensional scalar-product space $\Re^j$ labeled by the value of the angular momentum j. This space is "spanned" by the $2j+1$ eigenvectors $|j, m\rangle$ of the set of operators $\mathbf{J^2}$ and J_3 of (2.3.35). The space $\Re^j$ is the space in which all vectors are of the form

$$\Re^j = \{\psi^j = \sum_{m=-j}^{j} |j, m\rangle\, a_m, \quad a_m = \langle j, m|\psi^j\rangle \in \mathbb{C}\,\}\,, \tag{2.3.42}$$

where j and m are the eigenvalues of $\mathbf{J^2}$ and J_3. From (2.3.35), (2.3.38), and (2.3.41), the action of the operators $\mathbf{J^2}$, J_3, and $J_\pm$ on the basis vectors $|j, m\rangle$ is given by

$$\mathbf{J^2}|j, m\rangle = \hbar^2 j(j+1)\,|j, m\rangle\,, \tag{2.3.43a}$$

$$J_3|j, m\rangle = \hbar m\,|j, m\rangle\,, \tag{2.3.43b}$$

$$J_\pm|j, m\rangle = \hbar\sqrt{(j \pm m)(j \mp m + 1)}\,|j, m \pm 1\rangle\,. \tag{2.3.43c}$$

The action of any operator $A \in \mathscr{E}(\mathrm{SO}(3))$ can now be calculated by expressing it in terms of J_3, J_+, and J_- and then using (2.3.43).

The eigenvector $|j, m\rangle$ spans a one-dimensional space

$$\Re^j_m = \{\psi^j_m = a|j, m\rangle;\ \ a \in \mathbb{C}\,\}\,. \tag{2.3.44}$$

The projection operator onto this one-dimensional space,

$$\Lambda^j_m = |jm\rangle\langle jm|\,, \tag{2.3.45}$$

describes the state in which the angular momentum value is j and the component of angular momentum in the three-direction is m.

The space $\mathfrak{R}^j$ of (2.3.42) is said to be the orthogonal, direct sum of the one-dimensional spaces $\mathfrak{R}^j_m$ and is written as

$$\mathfrak{R}^j = \sum_{m=-j}^{j} \oplus \mathfrak{R}^j_m \,, \qquad \mathfrak{R}^j_{m'} \perp \mathfrak{R}^j_m \text{ for } m' \neq m \,. \tag{2.3.46}$$

The subspaces $\mathfrak{R}^j_m$ and $\mathfrak{R}^j_{m'}$ are orthogonal to each other when $m \neq m'$. Equation (2.3.42) implies that the $2j+1$ angular momentum eigenvectors

$$|j, m\rangle, \qquad m = j, j-1, j-2, \ldots, -j+1, -j \,, \tag{2.3.47}$$

are a basis system for the $(2j+1)$-dimensional space $\mathfrak{R}^j$ - just as the three vectors $\mathbf{e}_2$, $\mathbf{e}_2$, and $\mathbf{e}_3$ provide a basis system for the usual three-dimensional space $\mathbb{R}_3$. A major mathematical difference between $\mathbb{R}_3$ and $\mathfrak{R}^j$ is that in $\mathbb{R}_3$ the components x^i of the vector $\mathbf{x} = \mathbf{e}_i x^i$ are real numbers, $x^i \in \mathbb{R}$, and in $\mathfrak{R}^j$ the expansion coefficients a_m in (2.3.42) are complex, $a_m \in \mathbb{C}$.

The operators J_3, J_+, J_- (or J_1, J_2, J_3) and all elements A of the enveloping algebra of SO(3), $\mathscr{E}$(SO(3)), as defined in (2.3.8) act *irreducibly* in the space $\mathfrak{R}^j$. That is, for every $\psi^j \in \mathfrak{R}^j$ it follows that $A\,\psi^j = \tilde{\psi}^j \in \mathfrak{R}^j$. This means that there exists no non-trivial subspace that remains invariant (is not transformed) under the action of all the operators $A \in \mathscr{E}$(SO(3)). The space $\mathfrak{R}^j$ is called an irreducible representation space of $\mathscr{E}$(SO(3)).[4] In summary, for every integer or half integer value of j, there exists a $(2j+1)$-dimensional, linear, scalar-product space $\mathfrak{R}^j$ in which the angular momentum operators J_κ, ($\kappa = 0, \pm 1$), act as given in (2.3.43).

The space $\mathfrak{R}^j$ describes a quantum system with angular momentum j. Similarly, the space $\mathfrak{R}^j_m$ describes a quantum system with angular momentum j and component of angular momentum $J_3 = \mathbf{e}_3 \cdot \mathbf{J}$ in the three-direction $\mathbf{e}_3$ given by m. Here $\mathbf{e}_3$ is the unit vector in the 3-direction, which can be chosen to be in any arbitrary direction. When a specific direction is singled out, for example, by the direction of a magnetic field, then $\mathbf{e}_3$ is chosen as the direction of the magnetic field.[5] The

---

[4] As will be discussed below, $\mathfrak{R}^j$ is also the irreducible representation space of the set of all rotations in ordinary three-dimensional space $\mathbb{R}_3$. This set of rotations forms a group. A group G is a set of elements a, b, c,..., with the following properties:

(1) For every $a, b \in G$ then $a \cdot b \in G$ and $b \cdot a \in G$.

(2) For every $a, b, c \in G$, then $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(3) There exists an $e \in G$, called the identity, such that $e \cdot a = a \cdot e = a$ for all $a \in G$.

(4) For every $a \in G$ there exists exactly one $h \in G$ such that $h \cdot a = a \cdot h = e$. The element h is called the inverse of a and is written $h = a^{-1}$.

In general, $a \cdot b \neq b \cdot a$ so the elements do not necessarily commute.

[5] Since the magnetic field was used in early quantum mechanics to distinguish a particular direction in three-dimensional space, the quantum number m is called the magnetic quantum number.

eigenvectors of the component of angular momentum in any direction $\mathbf{n}$, namely $\mathbf{n} \cdot \mathbf{J}$, can be given as linear combinations of the form (2.3.42).

When the possible values of m of an irreducible representation are plotted along a line as in Fig. 2.3 on page 80: the figure is called the weight diagram of the irreducible representation characterized by j. To each dot there corresponds a basis vector $|j, m\rangle$ (up to a phase factor) in the representation space $\mathfrak{R}^j$. That is, there is a one-to-one correspondence between each dot and the one-dimensional subspace $\mathfrak{R}^j_m$. Each one-dimensional subspace represents a (pure) physical state with angular momentum j and component m. Thus to each point in the weight diagram there corresponds a (pure) physical state with angular momentum j and component of angular momentum m. This subspace is also described by the projection operator Λ^j_m of (2.3.45).

In general, angular momentum is defined by the commutation relations (2.3.1). As far as the algebra of angular momentum is concerned, there is no difference between the representations with integer and half-integer values of j. However, angular momentum is not always defined solely by the commutation relations (2.3.1). For example, orbital angular momentum L_i is defined in terms of the position operators Q_j and momentum operators P_k according to the relation

$$L_i \equiv \epsilon_{ijk} Q_j P_k \,, \qquad (\mathbf{L} = \mathbf{Q} \times \mathbf{P}) \,, \tag{2.3.48}$$

where Q_j and P_k fulfill the canonical commutation relations (2.2.6). The operator $\mathbf{L} = \mathbf{Q} \times \mathbf{P}$ represents orbital angular momentum of a mass-point, of the center of mass of an extended object, or of an arrangement of mass points such as the dumbbell.

As a result of the commutation relations satisfied by Q_j and P_k, it is straightforward to show that the L_i satisfy the algebra of angular momentum (2.3.1). But an additional consequence of the L_i being defined by (2.3.48) is that the eigenvalues of L_3 can only be integers, $m = 0, 1, 2, \ldots$, and the eigenvalues of $\mathbf{L}^2$ are $\ell(\ell + 1)$ with ℓ taking only integer values, $\ell = 0, 1, 2, \ldots$.[6]

Example 2.3.3 Show that the allowed values of ℓ for orbital angular momentum are integers.

Solution From (1.2.16) it follows that the annihilation and creation operators for the harmonic oscillator in three dimensions are, respectively,

$$a_i = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{\mu\,\omega}{\hbar}} Q_i + \frac{i}{\sqrt{\mu\,\omega\,\hbar}} P_i \right) \,, \tag{2.3.49a}$$

$$a_i^\dagger = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{\mu\,\omega}{\hbar}} Q_i - \frac{i}{\sqrt{\mu\,\omega\,\hbar}} P_i \right) \,. \tag{2.3.49b}$$

---

[6] When $\ell = 1$, the angular momentum operators L_i are the generators of rotations in three-dimensional space $\mathbb{R}_3$, which is isomorphic to $\mathfrak{R}^{\ell=1}$.

Solving for Q_i and P_i in terms of a_i and $a_i^\dagger$,

$$Q_i = \sqrt{\frac{\hbar}{2\,\mu\,\omega}}(a_i^\dagger + a_i)\,, \qquad P_i = i\sqrt{\frac{\mu\,\omega\,\hbar}{2}}(a_i^\dagger - a_i)\,. \tag{2.3.50}$$

From (2.2.3c)

$$L_3 = Q_1 P_2 - Q_2 P_1\,, \tag{2.3.51}$$

Using (2.3.49) to express Q_i and P_i in terms of raising and lowering operators,

$$L_3 = i\frac{\hbar}{2}[(a_1^\dagger + a_1)(a_2^\dagger - a_2) - (a_2^\dagger + a_2)(a_1^\dagger - a_1)] \tag{2.3.52}$$

From the Heisenberg commutation relations it follows that

$$[a_1^\dagger,\, a_2^\dagger] = 0\,, \qquad [a_1,\, a_2^\dagger] = 0\,, \qquad [\,a_1^\dagger,\, a_2] = 0\,, \qquad [a_1,\, a_2] = 0\,. \tag{2.3.53}$$

Combining (2.3.52) and (2.3.53), L_3 can be written in the form

$$L_3 = i\hbar\,(a_2^\dagger\, a_1 - a_1^\dagger\, a_2)\,. \tag{2.3.54}$$

To prove that all values ℓ of angular momentum are integers, it is convenient to introduce spherical components a_0 and $a_{\pm 1}$ of the a_i. In analogy to (2.3.21), the spherical components of the a_i are given by

$$a_0 \equiv a_3\,, \qquad a_{\pm 1} \equiv \mp\frac{1}{\sqrt{2}}(a_1 \pm i a_2)\,. \tag{2.3.55}$$

Inverting (2.3.55) to express the Cartesian components a_i in term of the spherical components,

$$a_1 = -\frac{1}{\sqrt{2}}(a_{+1} - a_{-1})\,, \qquad a_2 = \frac{i}{\sqrt{2}}(a_{+1} + a_{-1})\,, \qquad a_3 = a_0\,. \tag{2.3.56}$$

From the commutation relations (2.3.53), it follows that the spherical operators a_κ satisfy the commutation relations

$$[a_\kappa,\, a_\lambda^\dagger] = \delta_{\kappa,\lambda}\,, \qquad [a_\kappa,\, a_\lambda] = 0\,, \qquad [\,a_\kappa^\dagger,\, a_\lambda^\dagger] = 0\,, \tag{2.3.57}$$

where κ and λ independently take the values 0 and ± 1. Each of the three operators a_0, a_{+1} and a_{-1} obeys the algebra satisfied by the operator a as given in (1.3.17) that describes the one-dimensional harmonic oscillator. Also, for $\kappa \neq \lambda$, a_κ and $a_\kappa^\dagger$

commute with a_λ and $a_\lambda^\dagger$ so, just as is the case for the one-dimensional harmonic oscillator, it is possible to define three number operators

$$N_\kappa = a_\kappa^\dagger a_\kappa\,, \qquad \kappa = 0, \pm 1\,. \tag{2.3.58}$$

For each of the three spherical components there is a representation space $\mathscr{H}_\kappa$ with basis vectors $|n_\kappa\rangle$ that satisfy

$$N_\kappa |n_\kappa\rangle = n_\kappa |n_\kappa\rangle\,, \qquad n_\kappa = 0, 1, 2, \ldots\,. \tag{2.3.59}$$

In the direct-product space $\mathscr{H} = \mathscr{H}_0 \otimes \mathscr{H}_{+1} \otimes \mathscr{H}_{-1}$ the basis vectors are $|n_0\rangle \otimes |n_{+1}\rangle \otimes |n_{-1}\rangle$, and the number operator

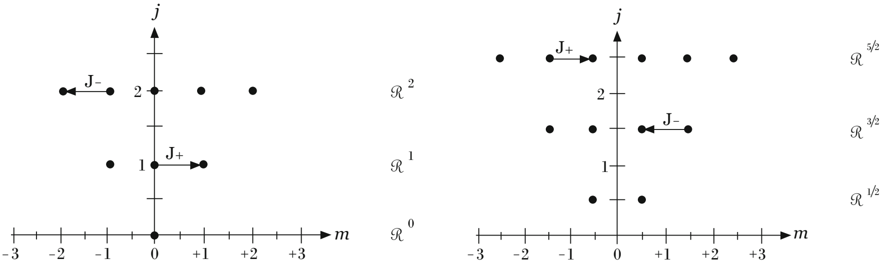
$$N = N_0 \otimes 1 \otimes 1 + 1 \otimes N_{+1} \otimes 1 + 1 \otimes 1 \otimes N_{-1}\,, \tag{2.3.60}$$

has eigenvalues $n = n_0 + n_{+1} + n_{-1}$ that are all integers. Using (2.3.56) to express L_3 as given in (2.3.54) in terms of spherical components a_κ (See Problem 2.11.),

$$L_3 = \hbar(a_{-1}^\dagger a_{-1} - a_{+1}^\dagger a_{+1}) = \hbar(N_{-1} - N_{+1})\,. \tag{2.3.61}$$

Since the eigenvalues of N_{-1} and N_{+1} are integers, all eigenvalues of L_3 are integers, implying that all values ℓ of orbital angular momentum are also integers.

In this section all solutions of the commutation relations for the angular momentum operators (2.3.1) have been derived. For every integer and half-integer value j, a solution has been shown to exist for the angular momentum commutation relations for self-adjoint, linear operators $J_i^{(j)}$ in the spaces $\Re^j$. (Usually the index ${}^{(j)}$ is omitted on the $J_i^{(j)}$). These $J_i^{(j)}$ in $\Re^j$ are called the irreducible representations of the angular momentum operators. For each representation associated with an eigenvalue j, there is a weight diagram similar to that shown in Fig. 2.3 on page 80. Weight diagrams for the lowest dimensional representations are characterized by $j = 0, 1/2, 1, \ldots$ and are given in Fig. 2.4. Each dot in the weight diagram



(a) (b)

Fig. 2.4 Weight diagrams for the irreducible representations of SO(3) with (**a**) $j = 0, 1, 2, \ldots$, and (**b**) $j = 1/2, 3/2, 5/2, \ldots$ and examples of the action of each of the operators J_+ and J_-

represents the space $\Re^j_m$, and the diagram also shows examples of the action of the operators $J_\pm$: $J_\pm \Re^j_m \to \Re^j_{m\pm1}$.

2.4 Observables of the Rotating Diatomic Molecule: Properties, Predictions and Limitations

2.4.1 The Algebra $\mathscr{E}(E(3))$ of Observables of the Rotating Dumbbell

When a quantum system is isolated and the only degrees of freedom are rotations, the only relevant observables are the angular momentum operators J_i, i = 1,2,3. The system will remain in a state described by the space $\Re^j$, where j is the angular momentum that the state happens to have at a particular time as the result of a prior preparation. A quantum system described by the $(2j+1)$-dimensional space $\Re^j$ is called an elementary rotator with angular momentum j (or $\hbar\, j$). The elementary rotator in a magnetic field will be studied in Chap. 5 when nuclear magnetic resonance is discussed. The rigid dumbbell, consisting of two mass points separated by a fixed distance, was initially discussed as a classical system in Sect. 2.2 and used to introduce angular momentum. The dumbbell is described by the classical position $\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_2$ that points along the rigid dumbbell axis as shown in Figs. 2.2 on page 70 and 2.5. Because the dumbbell is rigid, $\mathbf{r}^2$ = constant. Each mass point rotates about the center of mass, resulting in orbital angular momentum $\boldsymbol{\ell}$. If either or both of the mass points ("atoms") have intrinsic angular momentum $\mathbf{s}$ resulting from motion of the electrons or from electron spin, then the total angular momentum $\mathbf{j} = \boldsymbol{\ell} + \mathbf{s}$.

From this classical picture of the rotating, rigid dumbbell, it is conjectured that the corresponding *quantum mechanical rotating, rigid dumbbell* is described by six observables: three angular momentum operators J_i that correspond to the classical angular momenta j_i and three position operators Q_i that correspond to the three components r_i of the vector $\mathbf{r}$ that specifies the position of mass $m^{(1)}$ relative to mass $m^{(2)}$.

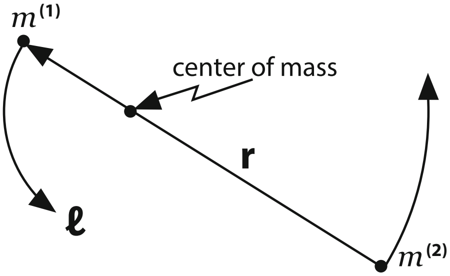


Fig. 2.5 Rigid dumbbell

2.4.2 Basic Observables—Fundamental Postulate III

For each physical system there exists a set of basic observables from which all others can be obtained as functions of the set of basic observables. When a quantum system has a classical analogue, the algebra of observables is obtained using the correspondence principle. For quantum systems with no classical analogue, the algebra of observables must be conjectured using intuition and trial-and-error.

Since the dumbbell is rigid, the operators Q_i satisfy the condition $\mathbf{Q}^2 = Q_i \cdot Q_i = r^2$ = constant. The quantum operators J_i , and Q_i correspond to the classical quantities j_i and r_i and fulfill the commutation relations

$$[J_i, J_j] = i\hbar\varepsilon_{ijk} J_k\,, \tag{2.4.1a}$$

$$[J_i, Q_j] = i\hbar\varepsilon_{ijk} Q_k\,, \tag{2.4.1b}$$

$$[Q_i, Q_j] = 0\,. \tag{2.4.1c}$$

The operators Q_i that fulfill the commutation relations (2.4.1b) with J_i, which are defined by the commutation relations (2.4.1a), are called *vector operators*.[7] The algebra of operators generated by the J_i and Q_i is denoted $\mathscr{E}$(E(3)) and is called the enveloping algebra of the Euclidean group E(3).[8]

The algebra of observables describing the rotating, rigid dumbbell is the algebra generated by the angular momentum operators J_i about the center of mass and the internal position operators Q_i describing the internuclear axis of the rotating, rigid dumbbell. Note that the position operators do not specify the position of the center of mass of the dumbbell. Such position operators would be required to describe the translational motion of the dumbbell as a whole, but that motion is being ignored here.

It can be shown (Problem 2.14) that the algebra of operators that satisfy (2.4.1) possess the property that the following set of operators commute with each other:

$$J_3\,, \qquad \mathbf{J}^2 = J_i J_i\,, \qquad \mathbf{Q}^2 = Q_i Q_i\,, \qquad \mathbf{Q}\cdot\mathbf{J} = Q_i J_i\,. \tag{2.4.2}$$

---

[7] Any set of operators F_i, $i = 1, 2, 3$, that fulfill the commutation relations $[J_i, F_j] = i\hbar\epsilon_{ijk} F_k$, where the J_i are the angular momentum operators, is called a vector operator.

[8] E(3) is the group of rotations SO(3) with generators J_i and the three position operators Q_i. In contrast, the physical translations in three-dimensional space are generated by the momentum operators P_i. The operators $m\,Q_i$ are the generators of the Galilean transformations. Thus the algebraic manipulations are based on group representation theory, a knowledge of which is not required here. The orbital angular momentum L_i defined by (2.2.3), and the commutation relations (2.4.1a) were initially derived from the commutation relations (2.2.6) satisfied by the P_i and Q_j. The orbital angular momentum number ℓ can take only integer values 0, 1, 2, … as proved in Sect. 2.3, Example 2.3.3 on page 83. The J_i are defined as self-adjoint operators fulfilling (2.4.1a) and are thus more general: the total angular momentum j can take both integer and half-integer values.

As a consequence, in the representation space $\mathfrak{R}$, in which the operators J_i and Q_i act, it is possible to choose vectors that are eigenvectors of all of the four operators in (2.4.2). These eigenvectors are denoted $|r, k_0, j, m\rangle$ where r, k_0, j and m label the eigenvalues of the operators in (2.4.2) as follows:

$$J_3|r, k_0, j, m\rangle = \hbar m|r, k_0, j, m\rangle\,, \tag{2.4.3a}$$

$$\mathbf{J}^2|r, k_0, j, m\rangle = \hbar^2 j(j+1)|r, k_0, j, m\rangle\,, \tag{2.4.3b}$$

$$\mathbf{Q}^2|r, k_0, j, m\rangle = r^2|r, k_0, j, m\rangle\,, \tag{2.4.3c}$$

$$\mathbf{Q}\cdot\mathbf{J}|r, k_0, j, m\rangle = \hbar r k_0|r, k_0, j, m\rangle\,. \tag{2.4.3d}$$

The system of operators (2.4.2) is called a complete system of commuting operators (c.s.c.o.) of the rotating, rigid dumbbell, and the set of quantum numbers $\{r, k_0, j, m\}$ is called a complete set of quantum numbers because any state of the dumbbell can be specified by these numbers.

From correspondence with the classical dumbbell depicted in Fig. 2.5 on page 86, it follows that the eigenvalue r^2 (or $r = \sqrt{r^2}$) characterizes the length of the internuclear, dumbbell axis. Similarly, the operator $\mathbf{Q}\cdot\mathbf{J}$ corresponds to the classical quantity $\mathbf{r}\cdot\mathbf{j}$, the scalar product of the vector $\mathbf{r}$ with the angular momentum $\mathbf{j}$. As a consequence, the eigenvalue $r\hbar\, k_0$ of the operator $\mathbf{Q}\cdot\mathbf{J}$ is the eigenvalue of the component of the angular momentum along the internuclear axis multiplied by the length of the internuclear axis, and $\hbar\, k_0$ is the component of angular momentum along the internuclear axis of the dumbbell molecule.

The operators $\mathbf{Q}^2$ and $\mathbf{Q}\cdot\mathbf{J}$ commute with all of the generators J_i and Q_i of the algebra of observables $\mathscr{E}(\mathrm{E}(3))$,

$$[\mathbf{Q}^2, Q_i] = 0\,, \qquad [\mathbf{Q}\cdot\mathbf{J}, Q_i] = 0\,, \qquad [\mathbf{Q}^2, J_i] = 0\,, \qquad [\mathbf{Q}\cdot\mathbf{J}, J_i] = 0\,. \tag{2.4.4}$$

Operators that commute with all generators of the algebra are called invariant operators of the algebra; therefore, $\mathbf{Q}^2$ and $\mathbf{Q}\cdot\mathbf{J} = Q_1J_1 + Q_2J_2 + Q_3J_3$ are invariant operators of the algebra $\mathscr{E}(\mathrm{E}(3))$.

Because $\mathbf{Q}^2$ and $\mathbf{Q}\cdot\mathbf{J}$ commute with all generators Q_i and J_i of the algebra $\mathscr{E}(\mathrm{E}(3))$, it follows from (2.4.3c) and (2.4.3d) that no operator J_i or Q_i can transform from a vector $|r, k_0, j, m\rangle$ to another vector $|r', k_0', j, m\rangle$ with eigenvalues $r' \neq r$ or $k_0' \neq k_0$. This is consistent with the previously imposed condition that the dumbbell is rigid, implying that $\mathbf{r}^2$ can only take one, fixed value and that the component k_0 of angular momentum along the internuclear axis, which is in the direction of $\mathbf{Q}$, is fixed. As a result, the operators J_i and Q_i act only within the single space $\mathfrak{R}(k_0, r)$ characterized by the eigenvalues (k_0, r), just as the angular momentum operators J_i act only within the single space $\mathfrak{R}^j$ characterized by the eigenvalue j. The pair of constants (k_0, r) labels the different ("inequivalent") representation spaces in which the generators J_i and Q_i of the algebra $\mathscr{E}(\mathrm{E}(3))$ act.

From (2.3.34), for every value j, the magnetic quantum number m takes the values $j, j-1, \cdots -j$. Thus the vectors $|r, k_0, j, m\rangle$, where (k_0, r) and j are fixed constants, span the $(2j+1)$-dimensional representation space of angular momentum $\Re^j$. In the general case j can be either an integer or a half-integer, but if the J_i are the orbital angular momentum operators (2.2.5) instead of the general angular momentum operators J_i, the values of j can only be integers.

Since $[J_i, \mathbf{Q} \cdot \mathbf{J}]=0$, $[J_i, \mathbf{Q}^2] = 0$, and $[J_i, \mathbf{J}^2]=0$, the operators J_i, and, more specifically, the raising and lowering operators $J_\pm = J_1 \pm iJ_2$, do not change the values (k_0, r) and j. The action of J_i on the basis vectors $|r, k_0, j, m\rangle$ changes only the value m. From (2.3.38) and (2.3.41),

$$J_\mp |r, k_0, j, m\rangle = \hbar\sqrt{(j \pm m)(j \mp m + 1)}|r, k_0, j, m \mp 1\rangle \,. \tag{2.4.5}$$

Similarly, from (2.3.43a),

$$J_3|r, k_0, j, m\rangle = \hbar m|r, k_0, j, m\rangle \,. \tag{2.4.6}$$

The operators J_i transform within an irreducible representation space $\Re^j$ with a specific value of angular momentum j, transforming between the one-dimensional subspaces with different angular momentum components m:

$$J_\mp \Re^j_m \longrightarrow \Re^j_{m\mp 1} \qquad J_3 \Re^j_m \longrightarrow \Re^j_m \,. \tag{2.4.7}$$

The action of the operators Q_i on the eigenvectors $|r, k_0, j, m\rangle$ can be calculated using techniques that are similar in principle, but much more complicated in detail, than those required to calculate $J_\pm|j, m\rangle$ as given in (2.3.38) and (2.3.41). Here the results of this purely mathematical calculation are given without proof.[9] Instead of using the Cartesian components Q_i, $i = 1, 2, 3$, it is again convenient to use the spherical components Q_κ, $\kappa = -1, 0, 1$, for the vector operator $\mathbf{Q}$ that are, in analogy with (2.3.21), defined by[10]

$$Q_0 \equiv Q_3 \,, \qquad Q_{\pm 1} = \mp\frac{1}{\sqrt{2}}(Q_1 \pm iQ_2) \,. \tag{2.4.8}$$

From (2.4.8) and the fact that $Q_i = Q_i^\dagger$, it immediately follows that

$$Q_0^\dagger = Q_0 \,, \tag{2.4.9a}$$

$$Q_{+1}^\dagger = -\frac{1}{\sqrt{2}}(Q_1 - iQ_2) = -Q_{-1} \,, \tag{2.4.9b}$$

---

[9] For more details and derivations see Chap. 5, Sect. 3 and the appendix to Chap. 5, Sect. 3 of A. Bohm, *Quantum Mechanics: Foundations and Applications*, Springer-Verlag, New York, 2nd Edition (1986), 3rd Edition (1993), and any later paperback or low-currency edition.

[10] Spherical components are defined in such a way that their matrix elements are readily expressed in terms of tabulated Clebsch-Gordan coefficients that will be discussed in Chap. 3, Sect. 3.5.

$$Q_{-1}^{\dagger} = \frac{1}{\sqrt{2}}(Q_1 + iQ_2) = -Q_{+1}\,. \tag{2.4.9c}$$

As previously discussed, when any operator acts on $|r, k_0, j, m\rangle$, the parameters r and k_0 remain unchanged. Because Q_κ is a vector operator, when Q_κ acts on $|r, k_0, j, m\rangle$, the change in the value of m is known. It remains to determine how Q_κ changes j. The result depends on the values of the parameters (k_0, r), and the derivation is given in the reference 9 on page 89.

$$\begin{aligned} Q_0|r, k_0, j, m\rangle = \sqrt{j^2 - m^2}\,C_j|r, k_0, j-1, m\rangle - ma_j|r, k_0, j, m\rangle \\ - \sqrt{(j+1)^2 - m^2}\,C_{j+1}|r, k_0, j+1, m\rangle\,, \end{aligned} \tag{2.4.10a}$$

$$\begin{aligned} -\sqrt{2}\;Q_{\pm 1}|r, k_0, j, m\rangle = \sqrt{(j \mp m)(j \mp m - 1)}\,C_j|r, k_0, j-1, m \pm 1\rangle \\ \mp \sqrt{(j \mp m)(j \pm m + 1)}\,a_j|r, k_0, j, m \pm 1\rangle \\ + \sqrt{(j \pm m + 1)(j \pm m + 2)}\,C_{j+1}|r, k_0, j+1, m \pm 1\rangle\,, \end{aligned} \tag{2.4.10b}$$

where

$$C_j = r\,\frac{i}{j}\sqrt{\frac{j^2 - k_0^2}{4j^2 - 1}} \quad \text{and} \quad a_j = \frac{k_0 r}{j(j+1)}\,. \tag{2.4.11}$$

The parameters (k_0, r) characterize the representations of the algebra of $\mathscr{E}(\mathrm{E}(3))$ and take the values,

$$k_0 = \pm 0, \pm 1/2, \pm 1, \pm 3/2, \pm \cdots \text{ and } 0 \le r < \infty\,. \tag{2.4.12}$$

Although (2.4.10) and (2.4.11) are not derived here, it is straight forward to verify some features of the equations. For example, because $\mathbf{J}^2$ does not commute with Q_j, Q_j changes the eigenvalue $j(j+1)$ of $\mathbf{J}^2$.

Example 2.4.1 Derive the equation

$$[\mathbf{J}^2, Q_j] = 2\hbar^2 Q_j + 2i\hbar\epsilon_{jki} Q_k J_i\,.$$

Solution Recalling that repeated indices are summed over their range and using the identity $[AB, C] = A[B, C] + [A, C]B$ that is readily verified by explicitly writing the commutators,

$$[\mathbf{J}^2, Q_j] \equiv [J_i J_i, Q_j] = J_i[J_i, Q_j] + [J_i, Q_j]J_i.$$

With the aid of (2.4.1b),

$$[\mathbf{J}^2, Q_j] = J_i(i\hbar\,\epsilon_{ijk} Q_k) + (i\hbar\,\epsilon_{ijk} Q_k) J_i = i\hbar\,\epsilon_{ijk}[J_i Q_k + Q_k J_i]\,.$$

Again using (2.4.1b), this time to commute J_i and Q_k,

$$[\mathbf{J}^2, Q_j] = i\hbar\,\epsilon_{ijk}[Q_k J_i + i\hbar\epsilon_{ikm} Q_m + Q_k J_i] = \hbar^2 \epsilon_{ikj}\epsilon_{ikm} Q_m + 2i\hbar\,\epsilon_{ijk} Q_k J_i\,.$$

The identity $\epsilon_{ikj}\epsilon_{ikm} = 2\delta_{jm}$ (See Problem 2.17.) allows the above equation to be written in the desired form.

Example 2.4.2 Determine the angular momentum of the vector $Q_{+1}|r, k_0, j, m = j\rangle$.

Solution From Example 2.4.1 and the definition (2.4.8) for $Q_{\pm 1}$, it follows that

$$[\mathbf{J}^2, Q_{\pm 1}] = 2\hbar^2 Q_{\pm 1} \pm 2\hbar Q_{\pm 1} J_3 + \sqrt{2}\hbar Q_3 J_{\pm}\,.$$

Applying the above equation to $|r, k_0, j, m = j\rangle$

$$\mathbf{J}^2 Q_{+1}|r, k_0, j, j\rangle = [Q_{+1}\mathbf{J}^2 + 2\hbar^2 Q_{+1} + 2\hbar Q_{+1} J_3 + \sqrt{2}\hbar Q_3 J_{+}|r, k_0, j, j\rangle\,,$$

When the raising operator J_+ acts on $|r, k_0, j, j\rangle$, the result is zero. Thus

$$\begin{aligned}\mathbf{J}^2 Q_{+1}|r, k_0, j, j\rangle &= \hbar^2\,[(j)(j+1) + 2 + 2j]\,Q_{+1}|r, k_0, j, j\rangle \\ &= \hbar^2 (j+1)(j+2) Q_{+1}|r, k_0, j, j\rangle\,,\end{aligned}$$

From the above equation it follows that $Q_{+1}|r, k_0, j, m = j\rangle$ is a state with angular momentum $j + 1$, which agrees with (2.4.10b).

Summarizing, for every pair of values (k_0, r) from the set (2.4.12), there exists a representation of $\mathscr{E}(\mathrm{E}(3))$ in a linear space $\Re(k_0, r)$ in which the eigenvectors $|r, k_0, j, m\rangle$ of the observables (2.4.2) form a basis system on which J_3, $J_\pm$, Q_0 and $Q_{\pm 1}$ act, respectively, as given by (2.3.43) and (2.4.10). As can be seen from (2.4.10), the spherical components Q_0 and $Q_{\pm 1}$ of the vector operator $\mathbf{Q}$ transform between neighboring angular momentum states, changing the value of j by +1, 0, and -1. Simultaneously Q_κ changes the value m by $\kappa = 0, \pm 1$:

$$Q_\kappa : j \begin{array}{l} \nearrow\; j+1 \\ \rightarrow\; j \\ \searrow\; j-1 \end{array} \qquad\qquad Q_\kappa : m \rightarrow m + \kappa. \tag{2.4.13}$$

The space $\Re(k_0, r)$ spanned by the eigenvectors $|r, k_0, j, m\rangle$ is the space of all linear combinations of $|r, k_0, j, m\rangle$:

$$\Re(k_0, r) = \{\psi^{(k_0,r)} = \sum_{j=k_0,\, k_0+1,\, \ldots}^{\infty} \sum_{m=-j,\, -j+1,\, \ldots}^{j} a_{j,m}^{(k_0,r)} |r, k_0, j, m\rangle, \qquad a_{j,m}^{(k_0,r)} \in \mathbb{C}\} \tag{2.4.14}$$

The Q_i and J_i are the generators of the algebra $\mathscr{A}(Q_i, J_i) = \mathscr{E}(\mathrm{E}(3))$, and the space $\Re(k_0, r)$ is transformed into itself by this algebra. $\Re(k_0, r)$ is an irreducible representation space of the algebra $\mathscr{A}(Q_i, J_i) = \mathscr{E}(\mathrm{E}(3))$.[11] The spaces $\Re(k_0, r)$ are invariant under the action of the elements of $\mathscr{A}(Q_i, J_i)$ in the same way that the spaces $\Re^j$ are invariant under the action of the J_i.

The eigenvalues (k_0, r) characterize the physical system described by the algebra of observables $\mathscr{A}(Q_i, J_i)$ in the space $\Re(k_0, r)$ in the same way that the eigenvalue j characterizes the physical system described by the algebra of observables $\mathscr{A}(J_i)$ generated by the J_i in the space $\Re^j$.

From (2.4.3c), it follows that the invariant r is the length of the internuclear axis, the "distance" between the two atoms of the diatomic dumbbell.[12] The physical interpretation of k_0 follows from (2.4.3d) and (2.4.10). According to (2.4.3d),

$$\frac{\mathbf{Q} \cdot \mathbf{J}}{r} |r, k_0, j, m\rangle = \hbar k_0 |r, k_0, j, m\rangle\,, \tag{2.4.15}$$

revealing that k_0 is the value of angular momentum along the direction $\mathbf{r}/r$ of the internuclear axis. Each of the eigenvectors of the form $|r, k_0, j-1, m+\kappa\rangle$ that appear on the right-hand side of (2.4.10) are multiplied by the coefficient C_j. But as can be seen from (2.4.11), $C_j = 0$ for $j = k_0$ so Q_κ cannot transform to a state with a value of j less than $j = k_0$. Thus for the quantum physical system described by the algebra of J_i and Q_k in the space $\Re(k_0, r)$, the lowest possible value of angular momentum that the system can have is k_0.

If the angular momentum is only the orbital angular momentum of the rigid dumbbell depicted in Fig. 2.5 on page 86 and the two atoms have no spin, then $J_i = L_i = \epsilon_{ijk} Q_j P_k$, implying that the lowest value of angular momentum is $k_0 = 0$. If additional angular momentum such as spin (intrinsic angular momentum) is involved, then k_0 could be any integer or half-integer.

---

[11] The algebra of observables $\mathscr{A}(Q_i, J_i)$ is the set of all (infinite) sums and products of operators of the Q_i and J_i; in this particular case it is the enveloping algebra of the Euclidean group E(3) of which Q_i and J_i are the group generators. The operators Q_i and J_i are generators of the associative algebra of observables and are also the generators of the group E(3), meaning, for example, that the rotation by an angle ϕ about the three-axis is given by the unitary operator $e^{-i\hbar\phi J_3}$ acting in the space $\Re(k_0, r)$ of (2.4.14).

[12] More precisely, if $e\mathbf{Q}$ is the dipole operator, then r is the distance between the center of positive electric charge and the center of negative electric charge.

The "rigid dumbbell rotator" discussed here consists of two atoms connected by a massless, rigid rod. For $k_0 = 0$ either the atoms are spinless or else the spins of the two atoms (electrons and nuclei) cancel so that the angular momentum is only the orbital angular momentum of the dumbbell $L_i = \epsilon_{ijk} Q_j P_k$. When $k_0 = 0$ the values of angular momentum j of the dumbbell are $j = 0, 1, 2, \ldots$.

When $k_0 = 0$, then a_j in (2.4.11) is also zero, $a_j = 0$. Therefore, from (2.4.10a) and (2.4.10b) it follows that there are no dipole matrix elements with $\Delta j = 0$. The operator Q_κ does not transform between states with the same j: it transforms between states with $\Delta j = \pm 1$:

$$\begin{tikzcd}[row sep=small] & j+1 & \\ Q_\kappa : j \arrow[ur] \arrow[dr] & & Q_\kappa : m \to m + \kappa. \\ & j-1 & \end{tikzcd} \tag{2.4.16}$$

The space of physical states for this rigid dumbbell rotator is given by

$$\Re(k_0 = 0, r) = \sum_{j=0,1,2,}^{\infty} \oplus \Re^j , \quad \text{where} \quad \Re^j = \sum_{m=-j}^{j} \oplus \Re^j_m , \tag{2.4.17}$$

and $\Re^j$ is the representation space with angular momentum j.

The action of the position operators Q_κ on the basis vectors $|r, k_0 = 0, j, m\rangle$ is given in (2.4.10): the effect of $Q_{\pm 1}$ and Q_0 on the space $\Re^j_m$ can be written in simplified form as

$$Q_{\pm 1} \Re^j_m \longrightarrow \Re^{j-1}_{m \pm 1} \oplus \Re^{j+1}_{m \pm 1} , \tag{2.4.18a}$$

$$Q_0 \Re^j_m \longrightarrow \Re^{j-1}_m \oplus \Re^{j+1}_m . \tag{2.4.18b}$$

All three operators $Q_{\pm 1}$ and Q_0 change the value of j while the $Q_{\pm 1}$ change the value of m and Q_0 does not. Repeatedly applying $Q_{\pm 1}$ and Q_0 to the vectors of one $\Re^j$ with angular momentum j ultimately yields the space $\Re(k_0 = 0, r)$as given in (2.4.17).

A graphical depiction of the infinite-dimensional space $\Re(k_0 = 0, r)$ in (2.4.17) is identical with the collection of weight diagrams of SO(3) depicted in Fig. 2.6 on the next page. Each dot represents a basis vector $|j, m\rangle = |r, k_0 = 0, j, m\rangle$ or, equivalently, the subspace $\Re^j_m$. Thus to each dot there corresponds a state of the rotating, spinless dumbbell. Each horizontal row of dots is a weight diagram for SO(3) as shown in Fig. 2.3 on page 80 that represents all states with a specific angular momentum j. The collection of weight diagrams of SO(3) shows which representation spaces $\Re^j$ of angular momentum j are contained in the space $\Re(k_0 = 0, r)$ and which angular momentum values are possible for the rotating dumbbell molecule.

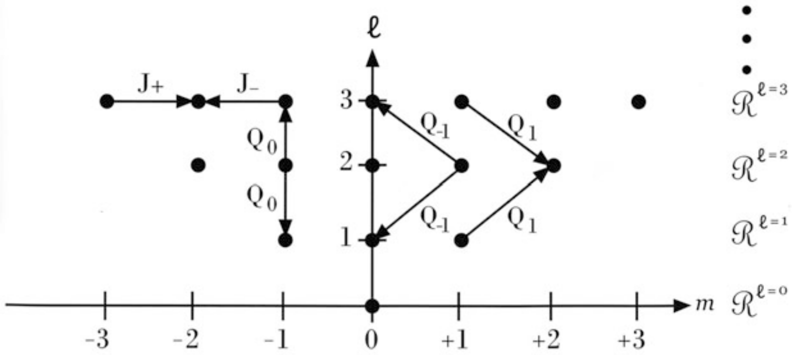


Fig. 2.6 Collections of weight diagrams of SO(3) that make up the infinite weight diagram of E(3) belonging to the representation space $\Re(k_0 = 0, r)$. Examples of the action of the operators $J_\pm$, Q_0 and $Q_{\pm 1}$ on the dots representing the basis vectors $|\ell, m\rangle$ are depicted by arrows, which show how the components of the dipole operator Q_κ, ($\kappa = 0, \pm 1$), transform between the energy levels of the rotating dumbbell molecule. Since the direct sum in (2.4.17) extends to infinity, the collections of dots extends to $\ell \to \infty$ as indicated by the dots above $\Re^{\ell=3}$

When $Q_{\pm 1}$ acts on the dot at the bottom of Fig. 2.6 that represents $\Re^{j=0}$, the space $\Re^{j=1}$ is obtained. The $J_\pm$ transform within each space $\Re^j$ as depicted in Fig. 2.4 on page 85. When $Q_{\pm 1}$ acts on the dots representing $\Re^{j=1}$, the spaces $\Re^{j=2}$ and $\Re^{j=0}$ can be obtained. Continuing in this way, "all" of the angular momentum spaces $\Re^j$, $j = 0, 1, 2, \ldots$ of the space of physical states of the diatomic molecule $\Re(k_0 = 0, r)$ are obtained by dipole transitions.[13] The weight diagram Fig. 2.6 of E(3) generated by J_i and Q_i extends to infinity since the space $\Re(k_0 = 0, r)$ is infinite dimensional.

From correspondence with the classical expression (2.2.30), the Hamiltonian H of the rigid, rotating diatomic molecule is conjectured to be given by

$$H = \frac{\mathbf{J}^2}{2I}, \tag{2.4.19}$$

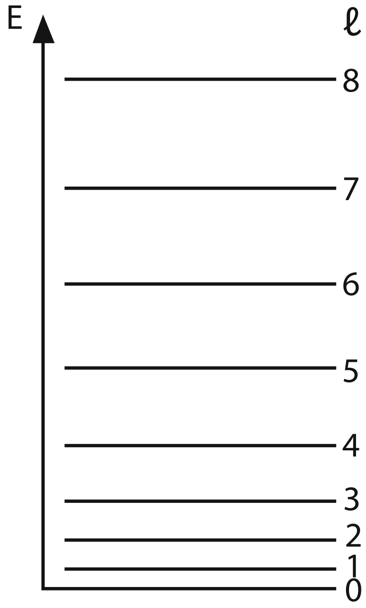
where I is the moment of inertia of the diatomic molecule. Applying this Hamiltonian to the vectors $|r, k_0 = 0, j, m\rangle$ in $\Re(k_0 = 0, r)$,

$$H|r, k_0 = 0, j, m\rangle = \frac{\mathbf{J}^2}{2I}|r, k_0 = 0, j, m\rangle \,. \tag{2.4.20}$$

---

[13] The spaces $\Re^j$ are the irreducible representation spaces of the rotation group SO(3) (or SU(2)). The space $\Re(k_0 = 0, r)$ for any real number r is also an irreducible representation space of a group, the three-dimensional Euclidean group E(3). The generators of this group are the J_κ and Q_κ, ($\kappa = 0, \pm 1$), or, equivalently, J_i and Q_i, $i = 1, 2, 3$, that obey the commutation relations (2.4.1) that are the defining commutation relations of the group E(3).

Fig. 2.7 Energy levels of a rigid, rotating diatomic molecule



Using (2.4.3b) yields the energy spectrum Fig. 2.7 of the rotator for which $k_0 = 0$. For such rotators the lowest energy eigenstate has angular momentum $\ell = 0$. The energy spectrum is given by

$$E_j = \frac{\hbar^2}{2I} j(j+1)\,, \qquad j = 0, 1, 2, \ldots\,. \tag{2.4.21}$$

To each energy value E_j there corresponds the (2j+1)-dimensional space $\Re^j$ defined in (2.3.46) that is described by the state operator

$$\rho^j = \frac{1}{2j+1} \sum_{m=-j}^{m=j} |j, m\rangle\langle j, m|\,. \tag{2.4.22}$$

Thus the energy value E_j is said to be $(2j + 1)$-fold degenerate. The energy levels (2.4.21) of the rigid, rotating diatomic molecules are depicted in Fig. 2.7. The energy levels (2.4.21) of the rigid rotator model agree with the experimentally observed energy levels of a rigid, rotating diatomic molecule in the limited energy range around 10^{-3} eV where the molecule is approximately rigid. At this energy the frequency of emitted and absorbed radiation is on the order of 10^{11} Hz.

The rotation group generated by the angular momentum operators J_i is a symmetry group. The J_κ, $(\kappa = 0, \pm 1)$, commute with the Hamiltonian H in (2.4.19) and do not transform between different energy values. The operators Q_κ transform between different energy levels (2.4.21) of the rigid rotator as indicated in (2.4.18) and given explicitly by (2.4.10). Therefore, the group E(3) with generators J_i and Q_i is not a symmetry group as is the rotation group SO(3) (or the Galilei group or Poincaré group of non-relativistic and relativistic space-time transformations, respectively). Instead E(3) describes the energy spectrum of the rotator and is, therefore, an example of the spectrum generating groups that were introduced about 1965.[14]

2.4.3 Absorption Spectrum for Diatomic Molecules

To obtain the spectrum of the radiation when the dumbbell rotator spontaneously emits photons, the following process must be considered:

$$\text{Dumbbell Rotator}^{\ell} \to \text{Dumbbell Rotator}^{j} + \gamma\,. \tag{2.4.23a}$$

Similarly, the absorption spectrum of the dumbbell rotator is determined by the process

$$\gamma + \text{Dumbbell Rotator}^{\ell} \to \text{Dumbbell Rotator}^{j}\,. \tag{2.4.23b}$$

The Einstein coefficients for the oscillator were introduced in Chap. 1, Sect. 1.5, and the Einstein coefficients A_{mn} for spontaneous emission and B_{mn} for induced absorption are given, respectively, by (1.5.19) and (1.5.31). In analogy to the discussion in Chap. 1, here the corresponding quantities will be determined for transitions between rotator energy levels. The discussion is slightly more complicated for rotator transitions because the rotator energy levels are degenerate: to one energy level there corresponds a $(2j+1)$-dimensional space instead of a single vector or a one-dimensional space.

The electric dipole moment for the rotator is the vector that points from the center of the negative charge distribution to the center of the positive charge distribution of the diatomic molecule. The centers of the positive and negative charge distributions do not coincide with the positions of the two masses of the dumbbell, but are instead determined by the *electron cloud of the molecule*. As a result of axial symmetry, the centers of the positive and negative charges lie on the internuclear axis. The electric dipole transitions are proportional to the square of the dipole moment $D_{nm} = q\,\langle n|Q|m\rangle$ where q is the magnitude of the electric charge of each of the two charges that form the dipole. If the molecules consist of two like atoms, as is the case for O_2, the centers of positive and negative charge distributions coincide, and there are no dipole transitions. But for molecules such as CO, consisting of unlike atoms, the electric dipole moment is nonzero, and dipole transitions occur.

---

[14] A. Bohm, Y. Neeman, A. O. Barut et.al. *Dynamical Groups and Spectrum Generating Algebras*, World Scientific Publishing Co., 1988.

In comparing the one-dimensional oscillator and the rotator in three dimensions, there are analogous quantities: The energy eigenspaces of the oscillator are one-dimensional and are described by the energy eigenvectors $|E_n\rangle$ of the oscillator. For the rotator the energy eigenspaces $\mathfrak{R}^j$ are, according to (2.3.42), (2j+1)-dimensional and are spanned by the eigenvectors $|r, k_0, j, m\rangle$. The parameters r and k_0 are not changed by the action of the J_κ, $(\kappa = 0, \pm 1)$, and the Q_κ because of (2.4.4). The J_κ do not change the value of j, but $J_\pm$ change the values of m as indicated in Fig. 2.6 on page 94. Also as indicated in Fig. 2.6 on page 94, the Q_κ change the values of j and m. The constants (r, k_0) characterize the specific diatomic molecule with a rigid dumbbell axis; therefore, they will often be suppressed in the discussion: $|r, k_0, j, m\rangle \to |j, m\rangle$.

The analog of the one-dimensional position operator Q for the oscillator are the position vector-operators Q^i, $i = 1, 2, 3$. Similarly, in place of the one-dimensional dipole matrix element D_{nm} of the oscillator, there is the vector dipole matrix element $D^i_{j\ell} = \langle j, j_3|Q_i|\ell, \ell_3\rangle$ for the rotator, where the indices j and ℓ, respectively, label the $(2j+1)$- and $(2\ell+1)$-dimensional energy eigenspaces $\mathfrak{R}^j$ and $\mathfrak{R}^\ell$ with energies E_j and E_ℓ, respectively.

The number $\mathscr{P}_{\rho'}(\Lambda^j)$ expresses the probability for the observable Λ^j in the state ρ' that has been transformed by the dipole operator $D_i = qQ_i$, $i = 1, 2, 3$, from the ℓth energy eigenstate ρ^ℓ.

The unpolarized energy state corresponding to the ℓth energy level E_ℓ in $\mathfrak{R}^\ell \equiv \Lambda^\ell \mathfrak{R}(k_0, r)$ is given by the state operator ρ^ℓ.

$$\rho^\ell = \frac{1}{2\ell+1}\Lambda^\ell \equiv \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} |\ell, m\rangle\langle\ell, m|\,, \tag{2.4.24}$$

where Λ^ℓ is the projection operator onto $\mathfrak{R}^\ell$.

To calculate dipole transitions, the action of the operator Q_i on the state ρ^ℓ must be considered. Since the position operators Q_i transform between different spaces $\mathfrak{R}^j$, the transitions generated by the position operators of the rotator are more general than for the oscillator. When the operators Q_i (or D_i) act on the state ρ^ℓ, the state that results is the transformed State ρ',

$$\rho' = \sum_i Q_i\rho^\ell Q_i = \frac{1}{2\ell+1}\sum_i Q_i\Lambda^\ell Q_i = \frac{1}{2\ell+1}\sum_i\sum_{m=-\ell}^{\ell} Q_i|\ell, m\rangle\langle\ell, m|Q_i\,. \tag{2.4.25}$$

The probability of a transition from the ℓth to the jth energy level is the probability of detecting the observable Λ^j in the transformed state ρ'. This probability, the Born probability for the observable Λ^j in the state ρ', is, according to axiom II of quantum mechanics, given by

$$\mathscr{P}_{\rho'}(\Lambda^j) = \mathrm{Tr}(\rho'\Lambda^j) = \sum_{i=1,2,3} \mathrm{Tr}[Q_i\rho^\ell Q_i\Lambda^j]. \tag{2.4.26}$$

For the oscillator the probability for $|E_n\rangle$ in the state ϕ' is denoted $\mathscr{P}_{\phi'}(|E_n\rangle) = |\langle E_n|\phi'\rangle|^2 = |\langle E_n|Q|E_m\rangle|^2$.

If all $(2j+1)$ polarizations are detected, the sum must be taken over all m' for the observable $\Lambda^j = \sum_{m'} |j,m'\rangle\langle j,m'|$. Using (2.4.25), the expression (2.4.26) for $\mathscr{P}_{\rho'}(\Lambda^j)$ becomes

$$\begin{aligned}\mathscr{P}_{\rho'}(\Lambda^j) &= \frac{1}{2\ell+1}\sum_i \sum_{m=-\ell}^{\ell} \mathrm{Tr}[Q_i|\ell,m\rangle\langle\ell,m|Q_i\Lambda^j] \\ &= \frac{1}{2\ell+1}\sum_i\sum_m\sum_{m'} \mathrm{Tr}[Q_i|\ell,m\rangle\langle\ell,m|Q_i|j,m'\rangle\langle j,m'|] \\ &= \frac{1}{2\ell+1}\sum_i\sum_m\sum_{m'} \langle j,m'|Q_i|\ell,m\rangle\langle\ell,m|Q_i|j,m'\rangle\,. \end{aligned} \tag{2.4.27}$$

$\mathscr{P}_{\rho'}(\Lambda^j)$ is the probability of detecting the observable Λ^j in the state ρ' that is obtained from the ℓth unpolarized energy eigenstate Λ^ℓ as a result of the action of the dipole operator Q_i. In other words, $\mathscr{P}_{\rho'}(\Lambda^j)$ is the probability for dipole transitions from the energy eigenspace $\mathfrak{R}^\ell$ to the energy eigenspace $\mathfrak{R}^j$. Because the $|\ell,m\rangle$ are vectors in the spherical basis, formulas are simpler when spherical components Q_κ, $(\kappa = 0, \pm1)$, of (2.4.8) are used instead of the Cartesian components Q_i of the vector transition operator. Making use of (2.4.8) and (2.4.9), it follows that

$$\mathbf{Q}^2 \equiv Q_iQ_i = Q_0^2 - Q_{+1}Q_{-1} - Q_{-1}Q_{+1} = Q_0^2 + Q_{-1}^\dagger Q_{-1} + Q_{+1}^\dagger Q_{+1}\,. \tag{2.4.28}$$

Summing over i in (2.4.27), and then using (2.4.28),

$$\begin{aligned}\mathscr{P}_{\rho'}(\Lambda^j) &= \frac{1}{2\ell+1}\sum_m\sum_{m'}[\langle j,m'|Q_0|\ell,m\rangle\langle\ell,m|Q_0|j,m'\rangle \\ &+ \langle j,m'|Q_{-1}^\dagger|\ell,m\rangle\langle\ell,m|Q_{-1}|j,m'\rangle + \langle j,m'|Q_{+1}^\dagger|\ell,m\rangle\langle\ell,m|Q_{+1}|j,m'\rangle] \\ &= \frac{1}{2\ell+1}\sum_m\sum_{m'}[\langle\ell,m|Q_0|j,m'\rangle^*\langle\ell,m|Q_0|j,m'\rangle \\ &+ \langle\ell,m|Q_{-1}|j,m'\rangle^*\langle\ell,m|Q_{-1}|j,m'\rangle + \langle\ell,m|Q_{+1}|j,m'\rangle^*\langle\ell,m|Q_{+1}|j,m'\rangle] \\ &= \frac{1}{2\ell+1}\sum_{\kappa=0,\pm1}\sum_m\sum_{m'}[\langle\ell,m|Q_\kappa|j,m'\rangle^*\langle\ell,m|Q_\kappa|j,m'\rangle]\,. \end{aligned} \tag{2.4.29}$$

The square of the electric dipole moment vector for the transitions with initial state unpolarized and all $(2j+1)$ polarizations detected is then given by

$$\sum_{i=1,2,3} D^i_{j,\ell}(D^i_{j,\ell})^\dagger = |\mathbf{D}_{j\ell}|^2 = \frac{q^2}{2\ell+1}\sum_i\sum_m\sum_{m'}\langle j,m'|Q_i|\ell,m\rangle\langle\ell,m|Q_i|j,m'\rangle\,, \tag{2.4.30}$$

which is often also written as

$$|\mathbf{D}_{j\ell}|^2 = \frac{q^2}{2\ell+1}\sum_m\sum_{m'}|\langle j,m'|\mathbf{Q}|\ell,m\rangle|^2\,. \tag{2.4.31}$$

The formula $|\langle j,m'|\mathbf{Q}|\ell,m\rangle|^2$ requires a summation be made over the three coordinates of the position operator,

$$|\langle j,m'|\mathbf{Q}|\ell,m\rangle|^2 = \sum_{i=1}^{3}|\langle j,m'|Q_i|\ell,m\rangle|^2 \tag{2.4.32a}$$

Using the definitions of Q_0 and $Q_{\pm1}$ as given in (2.4.8), the above formula can be rewritten in the desired form,

$$|\langle j,m'|\mathbf{Q}|\ell,m\rangle|^2 = |\langle j,m'|Q_0|\ell,m\rangle|^2+|\langle j,m'|Q_{+1}|\ell,m\rangle|^2+|\langle j,m'|Q_{-1}|\ell,m\rangle|^2\,. \tag{2.4.32b}$$

With the aid of (2.4.32b), (2.4.31) becomes

$$|\mathbf{D}_{j\ell}|^2 = \frac{q^2}{2\ell+1}\sum_m\sum_{m'}\Big[|\langle j,m'|Q_0|\ell,m\rangle|^2 + |\langle j,m'|Q_{+1}|\ell,m\rangle|^2 + |\langle j,m'|Q_{-1}|\ell,m\rangle|^2\Big]\,. \tag{2.4.33}$$

The above formula indicates that the “initial states” $|\ell,m\rangle$ with angular momentum ℓ must be averaged over, and the final states $|j,m'\rangle$ with angular momentum j must be summed over.

Although a rotating diatomic molecule emits magnetic dipole radiation, the magnetic dipole probabilities are approximately 10^{-5} of the electric dipole probabilities.[15] Thus the radiated energy results almost entirely from electric dipole radiation. As a consequence, the decay probability per unit time for the *spontaneous transition* from a level with angular momentum j into the ℓth energy level of the

---

[15] G. Herzberg, *Spectra of Diatomic Molecules*, 2nd Ed. p.19, Van Nostrand Reinhold Company, New York, 1950.

rotator is the Einstein coefficient $A_{j\ell}$ that is obtained from (1.5.19) by making the replacement $q^2|\langle E_n|q|E_m\rangle| \to |\mathbf{D}_{j\ell}|^2$,

$$A_{j\ell} = \frac{(\omega_{j,\ell})^3}{3\pi\epsilon_0 c^3\hbar}|\mathbf{D}_{j\ell}|^2 \,. \tag{2.4.34}$$

The decay probability per unit time is called the *decay rate*, which is denoted $\Gamma_{j\ell}$. Since $A_{j\ell}$ is the probability of decay per unit time, the lifetime $\tau_{j\ell}$ of a level with angular momentum j as it decays into the ℓth energy level of the rotator is the reciprocal of $A_{j\ell}$,

$$\tau_{j\ell} = \frac{1}{A_{j\ell}} = \frac{3\pi\epsilon_0 c^3\hbar}{(\omega_{j,\ell})^3}\frac{1}{|\mathbf{D}_{j\ell}|^2} \,. \tag{2.4.35}$$

The angular frequency $\omega_{j,\ell}$ in (2.4.34) and (2.4.35) is related to the frequency of the emitted radiation $\nu_{j,\ell}$ by $\omega_{j,\ell} = 2\pi\nu_{j,\ell}$. From energy conservation in the spontaneous emission process (2.4.23a), it follows that

$$E_\ell = E_j + h\nu_{j,\ell} \,. \tag{2.4.36}$$

Similarly the Einstein coefficient $B_{\ell j}$ for induced absorption is obtained from (1.5.30) by making the replacement $q^2|\langle E_n|q|E_m\rangle| \to |\mathbf{D}_{j\ell}|^2$,

$$B_{j\ell} = \frac{\pi}{3\epsilon_0\hbar^2}|\mathbf{D}_{j\ell}|^2 \,. \tag{2.4.37}$$

Energy conservation for induced absorption (2.4.23b) implies

$$h\nu_{j,\ell} + E_\ell = E_j \,. \tag{2.4.38}$$

The selection rules for spontaneous emission and induced absorption are obtained from the nonzero values for the matrix elements in (2.4.34). As determined analytically from (2.4.10) and indicated graphically in Fig. 2.6 on page 94, the matrix elements of Q_0 and $Q_{\pm 1}$ satisfy

$$\langle j, m'|Q_{\pm 1}|\ell, m\rangle = 0 \text{ unless } m' = m \pm 1 \text{ and } \ell - j = \pm 1 \,, \tag{2.4.39a}$$

$$\langle j, m'|Q_0|\ell, m\rangle = 0 \text{ unless } m' = m \text{ and } \ell - j = \pm 1 \,. \tag{2.4.39b}$$

The experimental results for the process of radiation *absorption* (2.4.23b) are now considered in detail. From (2.4.39) it follows that for absorption $j = \ell + 1$ so (2.4.38) becomes

$$h\nu_{\ell+1,\ell} = E_{\ell+1} - E_\ell \,. \tag{2.4.40}$$

In molecular spectroscopy it is customary to give the frequency in wave number units cm^{-1} where ν (in cm^{-1}) $=\nu$ (in Hz) $/c$ (in cm/s). Using the expression for E_ℓ as given in (2.4.21), from (2.4.39) the absorbed frequencies in wave number units are

$$\begin{aligned}\nu_{\ell+1,\ell} &= \frac{E_{\ell+1}-E_\ell}{hc} = \frac{\frac{\hbar^2}{2I}(\ell+1)(\ell+2)-\frac{\hbar^2}{2I}(\ell)(\ell+1)}{2\pi\hbar c}\\ &= \frac{\hbar}{4\pi cI}2(\ell+1) \equiv B\,2(\ell+1), \qquad \ell = 0,1,2,\ldots, \end{aligned} \tag{2.4.41}$$

where B is defined by

$$B \equiv \frac{\hbar}{4\pi cI} = \frac{h}{8\pi^2 cI}\,. \tag{2.4.42}$$

According to (2.4.41) the differences between the successive frequencies are given by

$$\Delta\nu = \nu_{\ell+1,\ell} - \nu_{\ell,\ell-1} = 2B, \tag{2.4.43}$$

implying that the frequencies are expected to have equidistant spacings. The energy levels of the rotator and the frequencies $\nu_{\ell+1,\ell}$ absorbed in the transition from the ℓth to the jth $= (\ell+1)$st level are shown in Fig. 2.7 on page 95 and Table 2.3 on page 103, respectively.

The absorption spectrum of HCl in the far infrared has been measured, and the experimental results, which are given in the second column of Table 2.3 on page 103, can be used to check the extent to which the model of the HCl molecule as a quantum mechanical dumbbell is justified. The third column of Table 2.3 on page 103 displays the differences between the successive frequencies. The first eleven frequencies have roughly equal spacing and fitting them with (2.4.41) yields the following value for the constant B:

$$B_{\mathrm{HCl}} = \frac{h}{8\pi^2 cI_{\mathrm{HCl}}} \approx 10.35\ \mathrm{cm}^{-1}\,. \tag{2.4.44}$$

The fourth column of Table 2.3 on page 103 gives the values calculated from (2.4.41) with the value (2.4.44). Comparing the second and fourth columns, there is fairly good agreement between the observed and calculated values for the first eleven frequencies, but the agreement becomes worse as ℓ increases further.

From (2.4.44) the moment of inertia I_{HCl} can be calculated,

$$I_{\mathrm{HCl}} = \frac{h}{8\pi^2 cB_{\mathrm{HCl}}} = 2.705\cdot 10^{-47}\ \mathrm{kg\ m}^2. \tag{2.4.45}$$

Using (2.2.22) and (2.2.27), the length r of the internuclear axis can be expressed in terms of the masses m_{H}, m_{Cl} and the moment of inertia I_{HCl},

$$r_{\mathrm{HCl}} = \sqrt{\frac{I_{\mathrm{HCl}}(m_{\mathrm{H}} + m_{\mathrm{Cl}})}{m_{\mathrm{H}} m_{\mathrm{Cl}}}} \,. \tag{2.4.46}$$

The mass of an atom can be calculated by dividing the mass per mole by Avogadro's number, which yields $m_{\mathrm{H}} = 1.674 \times 10^{-27}\,\mathrm{kg}$ and $m_{\mathrm{Cl}} = 5.887 \times 10^{-26}\,\mathrm{kg}$. Using the value of I_{HCl} obtained from the infrared absorption spectrum as given by (2.4.45), the internuclear distance between the hydrogen and chlorine nuclei can be calculated: $r_{\mathrm{HCl}} = 1.289 \times 10^{-10}\,\mathrm{m}$, in good agreement with values obtained from other classical considerations.

2.4.4 Absorption Spectrum for Non-rigid Diatomic Molecules

As ℓ increases, deviations from the rigid rotator model for the diatomic molecule become apparent when entries predicted by (2.4.41) in the fourth column of Table 2.2 are compared with the observed values in the second column. At higher values of ℓ, the differences between successive frequencies are smaller than at lower values. The frequencies are no longer equally spaced as predicted by the rigid rotator formula (2.4.41). Instead there is a trend suggesting that (2.4.41) should be modified to obtain better agreement at higher values of ℓ. The final column of the Table 2.3 on the next page fits the data with the empirical formula

$$\nu_{\ell+1,\ell} = 2b(\ell + 1) - 4d(\ell + 1)^3 \,, \tag{2.4.47}$$

Table 2.2 Basis vectors, position operator, dipole matrix elements, energy eigenstates and probability for a dipole transition of the oscillator and corresponding quantities for the rotator

	Oscillator		Rotator
Basis vectors	$\vert n\rangle$	$\longrightarrow$	$\vert \ell, m\rangle$
Position operator	Q	$\longrightarrow$	Q^i
Dipole matrix elements	$D_{mn} = \langle E_m \vert Q \vert E_n\rangle$	$\longrightarrow$	$D^i_{j\ell} = q\langle j, j_3 \vert Q_i \vert \ell, \ell_3\rangle$
Energy eigenstates	$\rho^n = \vert E_n\rangle\langle E_n \vert$ $\phi_n = \vert E_n\rangle$	$\longrightarrow$	$\rho^\ell = \frac{1}{(2\ell+1)} \sum_{m=-\ell}^{m=\ell} \times$ $\vert \ell, m\rangle\langle \ell, m \vert$ $= \frac{1}{(2\ell+1)} \Lambda^\ell$
Transition probability	$\mathscr{P}_{Q\vert m\rangle}(\vert n\rangle\langle n\vert)$ $=\mathrm{Tr}([\vert n\rangle\langle n \vert Q \vert m\rangle\langle m \vert Q])$ $= \vert\langle n \vert Q \vert m\rangle\vert^2$ $\phi' = Q\phi_n$	$\longrightarrow$	$\mathscr{P}_{\rho'}(\Lambda^j)$ where $\rho' = Q_i \rho^\ell Q_i$, and $\rho^\ell = \frac{1}{(2\ell+1)} \Lambda^\ell$

Table 2.3 Absorption spectrum of HCl in the far infrare

ℓ	$\nu_{\ell-1,\ell}$ cm^{-1}	$\Delta\nu_{obs}$	$\nu_{calc} = 20.70\ell$	$\nu_{calc} = 20.88\ell - 0.001837\ell^3$
1	20.8	20.8	20.70	20.87
2	41.6	20.9	41.41	41.74
3	62.5	20.53	62.11	62.58
4	83.03	21.10	82.82	83.39
5	104.13	20.60	103.52	104.15
6	124.73	20.64	124.22	124.86
7	145.37	20.52	144.93	145.50
8	165.89	20.34	165.63	166.07
9	186.23	20.37	186.33	186.54
10	206.60	22.26	207.04	206.92
11	228.86		227.74	227.19
17	345.6	19.0	351.96	345.86
18	364.6	19.6	372.67	365.05
19	384.2	19.2	393.37	384.04
20	403.4	18.7	414.07	402.82
21	422.1	18.0	434.78	421.37
22	440.1	18.2	455.48	439.70
23	458.3	18.1	476.18	457.79
24	474.4	19.9	496.89	475.62
25	494.3	16.9	517.59	493.18
26	511.2	16.5	538.30	510.48
27	527.7	15.2	559.00	527.48
28	542.9	18.2	579.70	544.19
29	561.1	15.3	600.41	560.59
30	576.4	13.2	621.11	576.66
31	589.6	19.8	641.81	592.41
32	609.4	14.0	662.52	607.82
33	623.4		683.22	622.87

Data for $l = 1, 2, 3$ from McCubbin *J. Chem. Phys.* **20**, 668 (1952); for $l = 4, \ldots, 11$ from R. L. Hansler and R. A. Oetjen, *J. Chem. Phys.* **21**, 1340 (1953); for $l = 17, \ldots, 33$ from J. Strong, *Phys. Rev.* **45**, 877 (1934). The units of ν are cm^{-1}

where b and d are constants. Comparing the final column with the observed values in the second column, the agreement of (2.4.47) with the experimental data is far better than that of (2.4.41). The energy spectrum that corresponds to (2.4.47) is given by[16]

$$E_l = 2\pi\hbar c[b\ell(\ell+1) - d\ell^2(\ell+1)^2]\,. \tag{2.4.48}$$

---

[16] The *ad hoc* formula $E_\ell = 2\pi\hbar c[b' - d'\ell]\ell(\ell+1)$ is popular for fitting rotational spectra; however, (2.4.47) provides a better fit to the molecular rotation spectra. It will be shown that (2.4.47) is easily understood theoretically.

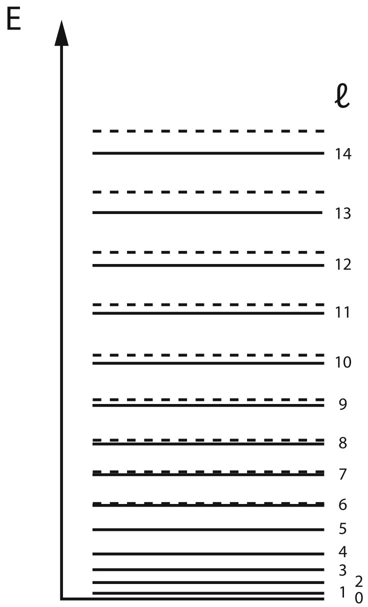


Fig. 2.8 Energy levels of the non-rigid rotator (solid lines) and the corresponding rigid rotator (dashed lines). To the scale of the drawing, the levels cannot be distinguished for $\ell < 6$ [from Herzberg (1966), with permission]

The energy levels (2.4.48) have been drawn in Fig. 2.8 with an exaggerated value of d that appears in the above equation.

The explanation for the better fit of (2.4.48) to the experimental values follows from the fact that the diatomic molecule HCl is not exactly a rigid rotator. The bonds between atoms are not rigid with the result that the interatomic distance increases with the speed of rotation. Equation (2.4.48) can be obtained theoretically by returning to the classical picture in which the molecule is considered to be two mass points (nuclei of atoms) joined, not by a rigid rod as in Fig. 2.2 on page 70 , but instead by a spring as in Fig. 2.9 on the facing page. This is the same picture used in Chap. 2 to describe the vibrating dumbbell of the quantum harmonic oscillator. The respective equilibrium distances between the center of mass and the masses $m^{(1)}$ and $m^{(2)}$ when the molecule is not spinning are denoted $\mathbf{x}_{\mathrm{e}}^{(1)}$ and $\mathbf{x}_{\mathrm{e}}^{(2)}$ and are depicted in Fig. 2.9 on the next page. The total distance between the masses when the molecule is spinning and when the molecule is not are, respectively, $\mathbf{r}$ and $\mathbf{r}_{\mathrm{e}}$ where $\mathbf{r}$ is given in (2.2.18) and $\mathbf{r}_{\mathrm{e}}$ is

$$\mathbf{r}_{\mathrm{e}} = \mathbf{x}_{\mathrm{e}}^{(1)} - \mathbf{x}_{\mathrm{e}}^{(2)} . \tag{2.4.49}$$

The energy of this system is the sum of the rotational kinetic energy of the two nuclei plus the (elastic) potential energy of the oscillator,

$$E = \frac{1}{2} m^{(1)} \left[\frac{\mathrm{d}\mathbf{x}^{(1)}}{\mathrm{d}t} \right]^2 + \frac{1}{2} m^{(2)} \left[\frac{\mathrm{d}\mathbf{x}^{(2)}}{\mathrm{d}t} \right]^2 + \frac{1}{2} k (r - r_{\mathrm{e}})^2 . \tag{2.4.50}$$

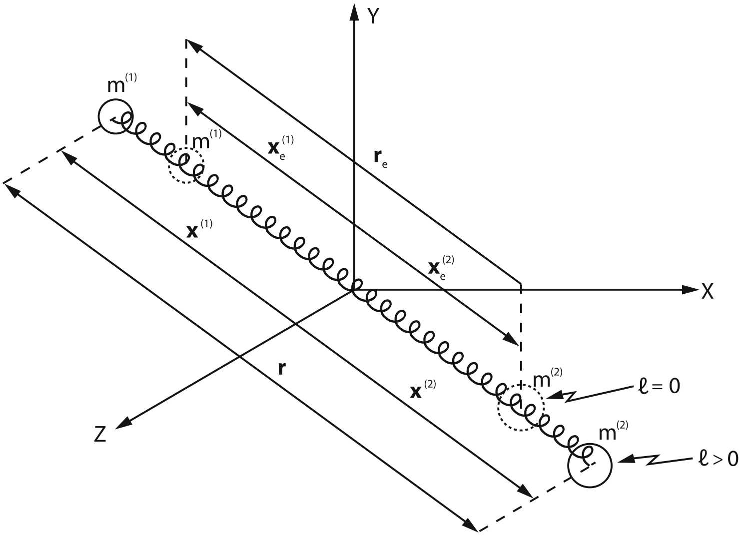


Fig. 2.9 For a rotating diatomic molecules with a variable internuclear distance, the position vectors $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ for masses $m^{(1)}$ and $m^{(2)}$, respectively, and the vector $\mathbf{r} = \mathbf{x}^{(1)} - \mathbf{x}^{(2)}$; the position vectors $\mathbf{x}_\mathrm{e}^{(1)}$ and $\mathbf{x}_\mathrm{e}^{(2)}$ for masses $m^{(1)}$ and $m^{(2)}$, respectively, when $\ell = 0$ and the vector $\mathbf{r}_\mathrm{e} = \mathbf{x}_\mathrm{e}^{(1)} - \mathbf{x}_\mathrm{e}^{(2)}$

In the above equation k is the "spring constant."

With the aid of (2.2.21) and (2.2.25), the two kinetic terms in (2.4.50) can be expressed in terms of the variables $\mathbf{l}^2$ and r,

$$E_\text{kinetic} = \frac{1}{2}m^{(1)}\left[\frac{\mathrm{d}\mathbf{x}^{(1)}}{\mathrm{d}t}\right]^2 + \frac{1}{2}m^{(2)}\left[\frac{\mathrm{d}\mathbf{x}^{(2)}}{\mathrm{d}t}\right]^2 = \frac{\mathbf{l}^2}{2\mu r^2}. \tag{2.4.51}$$

The elastic force keeps the two masses moving in a circle about the center of mass so, for each of the two masses, Newton's second law requires that the elastic force equal the product of the mass of the particle and its centripetal acceleration,

$$k(r - r_\mathrm{e}) = m_i \frac{\left(\frac{\mathrm{d}x^{(i)}}{\mathrm{d}t}\right)^2}{|x^{(i)}|}, \qquad i = 1, 2\,. \tag{2.4.52}$$

Using (2.2.20) with $\mathbf{x}_\text{CM} = 0$, the right-hand side of (2.4.52) is found to be the same for both $i = 1$ and $i = 2$, and the following expression for $r - r_\mathrm{e}$ is obtained:

$$r - r_\mathrm{e} = \frac{\mu}{kr}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 = \frac{\mathbf{l}^2}{k\mu r^3}\,. \tag{2.4.53}$$

The final equality follows from (2.2.25).

Equations (2.4.51) and (2.4.53) permit the formula for energy as given in (2.4.50) to be written in the form

$$E = \frac{\mathbf{l}^2}{2\mu r^2} + \frac{(\mathbf{l}^2)^2}{2k\mu^2 r^6} \,. \tag{2.4.54}$$

To determine the spectrum of the Hamiltonian corresponding to the expression for the energy E in (2.4.54), it is necessary to determine the variable r^2, which is accomplished by expressing r in terms of $\mathbf{l}^2$. Adding and subtracting r_e in the numerator of $(r/r_\mathrm{e})^2$,

$$\left(\frac{r}{r_\mathrm{e}}\right)^2 = \left[\frac{r_\mathrm{e} + (r - r_\mathrm{e})}{r_\mathrm{e}}\right]^2 = 1 + 2\frac{(r - r_\mathrm{e})}{r_\mathrm{e}} + O\left(\left[\frac{(r - r_\mathrm{e})}{r_\mathrm{e}}\right]^2\right) \,. \tag{2.4.55}$$

Using (2.4.53) to rewrite $r - r_\mathrm{e}$ in terms of $\mathbf{l}^2$,

$$r^2 = r_\mathrm{e}^2 + \frac{2r_\mathrm{e}}{k\mu r^3}\mathbf{l}^2 + O\left((\mathbf{l}^2)^2\right) \,. \tag{2.4.56}$$

The desired formula for energy is obtained using (2.4.56) to eliminate r^2 in (2.4.54),

$$E = \frac{\mathbf{l}^2}{2\mu(r_\mathrm{e}^2 + \frac{2r_\mathrm{e}}{k\mu r^3}\mathbf{l}^2)} + \frac{(\mathbf{l}^2)^2}{2k\mu^2 r_\mathrm{e}^6} + O\left((\mathbf{l}^2)^3 = \frac{\mathbf{l}^2}{2\mu r_\mathrm{e}^2} - \frac{(\mathbf{l}^2)^2}{2k\mu^2 r_\mathrm{e}^6} + O\left((\mathbf{l}^2)^3\right) \,. \tag{2.4.57}$$

The first term is the energy of the rigid rotator, and the second term is the elastic energy that is present because the rotator is not precisely rigid. The corresponding quantum Hamiltonian is obtained by replacing the number $\mathbf{l}^2$ with the operator $\mathbf{J}^2$,

$$H = \frac{\mathbf{J}^2}{2\mu r_\mathrm{e}^2} - \frac{(\mathbf{J}^2)^2}{2k\mu^2 r_\mathrm{e}^6} \,. \tag{2.4.58}$$

The energy levels of the above Hamiltonian are those given by (2.4.48) that in turn yield the spectrum (2.4.47) for the absorbed frequencies given in the fifth column of Table 2.3 on page 103. The very good fit that (2.4.47) provides to the experimental data confirms the validity of the above classical considerations. From the experimental data in the second column in Table 2.3 on page 103, the values obtained for the parameters b_HCl and d_HCl are

$$b_\mathrm{HCl} = 10.438 \text{ cm}^{-1}, \quad d_\mathrm{HCl} = 0.00046 \text{ cm}^{-1}. \tag{2.4.59}$$

The fact that d_HCl is orders of magnitude smaller than b_HCl verifies that the rigid rotator is a remarkably good model of the rotating diatomic molecule.

As discussed in this section, the rigid rotator provides a good description of the absorption spectra of a diatomic molecule. To obtain a more accurate description, however, corrections and modifications had to be incorporated, providing an even better understanding of the physics. Here the diatomic molecule has been described as a quantum rotator, and in Chap. 1 it was described as a quantum oscillator. The simple dumbbell model of a diatomic molecule as a non-rigid rotator that undergoes harmonic oscillations provides a good description of the experimental data. These models are benchmarks, and the physics of the real quantum systems is understood by the agreement with and deviations from these benchmarks.

2.5 Angular Momentum States and Calculation of Probabilities

According to Fundamental Postulate II of quantum mechanics (See Chap. 1, Sect. 1.4.), the state of a quantum system is described by a density or statistical operator. In the space where experiments take place, a state with angular momentum j for which any component m is equally probable is described by a state operator

$$\rho^j = (\mathrm{Tr}\,\Lambda^j)^{-1}\Lambda^j = \frac{1}{(2j+1)}\sum_{m'=-j}^{j} |j,m'\rangle\langle j,m'|\,. \tag{2.5.1}$$

Λ^j acts the unit operator in $\mathfrak{R}^j$ and is also the projection operator onto the space $\mathfrak{R}^j$ from a larger space $\mathfrak{R} \supset \mathfrak{R}^j$. The fact that any component m is equally probable for the state ρ^j is established as follows.

According to Fundamental Postulate II of quantum mechanics, the probability of an observable A in the state ρ^j is calculated as the Born probability

$$\mathscr{P}_{\rho^j}(A) = \mathrm{Tr}(A\rho^j)\,. \tag{2.5.2}$$

Choosing the observable A to be the operator $|j,m\rangle\langle j,m|$, which represents the property that the three-component of angular momentum is m, the probability of obtaining a specific value m for the component of J_3 along the arbitrary direction e_3 in the state ρ^j is (Problem 2.25)

$$\begin{aligned}\mathscr{P}_{\rho^j}(|j,\,m\rangle\langle j,\,m|) &= \mathrm{Tr}(|j,m\rangle\langle j,m|\rho^j) = \langle j,m|\rho^j|j,m\rangle\,,\\ &= \frac{1}{2j+1}\sum_{m'=-j}^{j}\langle j,m|j,m'\rangle\langle j,m'|j,m\rangle = \frac{1}{2j+1}\,.\end{aligned} \tag{2.5.3}$$

The above equation reveals that each value m for the three-component of angular momentum J_3 is equally likely in the state ρ^j of (2.5.1). If an arbitrary direction in space is denoted e_3, then according to (2.5.3), the values for the component $J_3 = e_3 \cdot \mathbf{J}$ are $m = -j, j-1, \ldots + j$, and each of these values will appear in a measurement of J_3 with the same probability $1/(2j+1)$.

If, for a specific quantum system, no physical quantity distinguishes a direction, then the quantum number m is irrelevant. All that can be said is that the quantum system is in a state ρ^j as given in (2.5.1) with angular momentum j in which any component of angular momentum m is equally probable.

The quantum mechanical state ρ^j describes an ensemble of individual microsystems, which could, for example, be a large number N of diatomic molecules such as HCl or CO. The number N should be a large so that statistical statements can be made: N could be on the order of Avogadro's number 6.02×10^{23} although in some cases it could be as small as 200 or even smaller. If a preferred direction $\mathbf{e}_3$ is distinguished, for example, by the direction of a magnetic field, then a measurement of the angular momentum along the direction $\mathbf{e}_3$ would find $N/(2j+1)$ diatomic molecules with angular momentum component $\hbar m$ for each value of $m = -j, -j + 1, \ldots, +j$.

If the value of J_3 is measured in the state ρ^j of (2.5.1), then according to Fundamental Postulate II, the expectation value for the observable J_3 is

$$\begin{aligned}\mathrm{Tr}(J_3\rho^j) &= \frac{1}{2j+1}\sum_{m=-j}^{j} \mathrm{Tr}(J_3|jm\rangle\langle jm|) \\ &= \frac{1}{2j+1}\sum_{m=-j}^{j} \langle jm|J_3|jm\rangle = \frac{1}{2j+1}\sum_{m=-j}^{j} m = 0\,. \qquad (2.5.4)\end{aligned}$$

The expectation value for the three-component J_3 of angular momentum is zero, not because all N molecules have a zero three-component, but because there are an equal number $N/(2j+1)$ of molecules with component m and with component $-m$, so that the average value is zero.

Since the observables J_i do not commute, their values measured in a pure state $|j, m\rangle\langle j, m|$ will not be known with certainty. The expectation value of the operator J_3 and J_3^2 in the eigenstates $|j, m\rangle$ of $\mathbf{J}^2$ and J_3 is easily obtained from (2.3.35) :

$$\langle j, m|\mathbf{J}^2|j, m\rangle = \hbar^2 j(j+1), \quad \langle j, m|J_3|j, m\rangle = \hbar m\,. \qquad (2.5.5)$$

The uncertainty of the observable A in the state ρ is

$$\Delta_\rho A = \sqrt{\mathrm{Tr}(\rho A^2) - (\mathrm{Tr}(\rho A))^2}\,. \qquad (2.5.6)$$

As a consequence, the uncertainty of the observable J_3 in the state $\rho = |j,m\rangle\langle j,m|$ is

$$\begin{aligned}\Delta_{|j,m\rangle} J_3 &= \sqrt{\mathrm{Tr}(|j,m\rangle\langle j,m|J_3^2) - (\mathrm{Tr}(|j,m\rangle\langle j,m|J_3))^2} \\ &= \sqrt{\langle j,m|J_3^2|j,m\rangle - \langle j,m|J_3|j,m\rangle^2} = \hbar\sqrt{m^2 - m^2} = 0\,. \end{aligned} \tag{2.5.7}$$

The above result is expected since $|j,m\rangle$ is an eigenstate of J_3.

The expectation values of J_1 and J_1^2 in the state $|j,m\rangle\langle j,m|$ follow immediately from Examples 2.3.1 and 2.3.2 on page 81, respectively, and are given below for $m = j$:

$$\langle j,j|J_1|j,j\rangle = 0, \quad \langle j,j|J_1^2|j,j\rangle = \frac{\hbar^2 j}{2} \tag{2.5.8a}$$

Corresponding formulas for J_2 are calculated similarly and are as follows:

$$\langle j,j|J_2|j,j\rangle = 0, \quad \langle j,j|J_2^2|j,j\rangle = \frac{\hbar^2 j}{2} \tag{2.5.8b}$$

Although the average values of J_1 and J_2 are zero in the eigenstate $|j,j\rangle$ of J_3 for which $m = j$, the average of the squares of their values $\langle jj|J_1^2|jj\rangle$ and $\langle jj|J_2^2|jj\rangle$ are not; therefore, the uncertainties $\Delta_{|j,j\rangle} J_1$ and $\Delta_{|j,j\rangle} J_2$) are not zero. The uncertainty for the operator J_1 in the state $|j,j\rangle$ is

$$\Delta_{|j,j\rangle} J_1 = \sqrt{\langle j,j|J_1^2|j,j\rangle - \langle j,j|J_1|j,j\rangle^2} = \hbar\sqrt{\frac{j}{2}}\,. \tag{2.5.9a}$$

Similarly,

$$\Delta_{|j,j\rangle} J_2 = \hbar\sqrt{\frac{j}{2}}\,. \tag{2.5.9b}$$

Equation (2.5.9a) reveals that in a measurement of J_1 in the pure state $|j,j\rangle$, the values for the component J_1, which is "orthogonal" to J_3, will not be zero with certainty. This result is not intuitive: In classical physics, if the vector $\mathbf{j} = j_1\,\mathbf{e}_1 + j_2\,\mathbf{e}_2 + j_3\,\mathbf{e}_3$ points in the $\mathbf{e}_3$-direction, then $j_1 = j_2 = 0$, and $|\mathbf{j}| = j_3$. Thus j_1^2 and j_2^2 are with certainty zero. In contrast (2.5.9) shows that there is "uncertainty" regarding the values of J_1 and J_2. This feature is a quantum effect and occurs because

$$\langle j,j|\mathbf{J}^2 - J_3^2|j,j\rangle = \hbar^2[j(j+1) - j^2] = \hbar^2\, j \neq 0. \tag{2.5.10}$$

It immediately follows from (2.5.10) that the expectation value of $J_1^2 + J_2^2$ must be nonzero:

$$\langle j, j|J_1^2 + J_2^2|j, j\rangle = \langle j, j|\mathbf{J}^2 - J_3^2|j, j\rangle = \hbar^2[j(j+1) - j^2] = \hbar^2 j \neq 0. \tag{2.5.11}$$

This feature, which is very strange classically, is connected with the uncertainty relation (Problem 2.26).

2.6 Relationship Between $SU(2)$ and $SO(3)$

For integer values of j, the spaces $\mathfrak{R}^j$ are representation spaces of the rotation group $SO(3)$; for half-integer j, the spaces $\mathfrak{R}^j$ are representation spaces of the group $SU(2)$, which is in a two-to-one correspondence with the rotation group: To every rotation $\mathfrak{R} \in SO(3)$ there correspond two elements U and $-U$ in $SU(2)$.

$$\begin{tikzcd} & U \arrow[dr, leftrightarrow] & \\ SU(2) \arrow[ur, phantom, sloped, "\ni"] \arrow[dr, phantom, sloped, "\ni"] & & \mathfrak{R} \in SO(3) \\ & -U \arrow[ur, leftrightarrow] & \end{tikzcd}$$

For quantum mechanics, where the observable quantities are the probabilities, this two-to-one correspondence does not make any difference. To illustrate this consider the case of an observable represented by the projection operator $|\psi\rangle\langle\psi|$ and a pure state represented by the vector ϕ. The probability for the observable $|\psi\rangle\langle\psi|$ in the state ϕ is given, not by the matrix element (ψ, ϕ), but instead by the absolute value of the matrix element:

$$\mathscr{P}_\phi(|\psi\rangle\langle\psi|) = \mathrm{Tr}\,(|\psi\rangle\langle\psi|\phi\rangle\langle\phi|) = |\langle\psi|\phi\rangle|^2\,. \tag{2.6.1}$$

In the spaces $\mathfrak{R}^j$, where j is a half-integer, for a rotation R there exist both $U(R)$ and $-U(R)$. For half-integer values of j, the rotated state of the state vector $\phi \in \mathfrak{R}^j$ is, therefore, given by $U\phi$ and $-U\phi$. But the probability of the observable ψ in the transformed state $U\phi$ or the transformed state $-U\phi$ is the same:

$$\mathscr{P}_{-U\phi} = |(\psi, -U\phi)|^2 = |-(\psi, U\phi)|^2 = |(\psi, U\phi)|^2 = \mathscr{P}_{U\phi} \tag{2.6.2}$$

The physical state, instead of being represented by the vector ϕ, is represented by the one-dimensional subspace spanned by the vector $\{\phi(\alpha) \;:\; \phi(\alpha) = e^{i\alpha}\phi\}$, where α is real, or, equivalently, by the projection operator $|\phi\rangle\langle\phi|$ onto this subspace.

Symmetry transformations such as the group of rotations of the apparatus in three-dimensional space are not described by representations of the rotation group, but instead by "representations up to a factor," which are also called "projective representations." The set of "projective representations" of the rotation group $SO(3)$ is the same as the set of representations of $SU(2)$, and the set of all unitary, irreducible representations of $SU(2)$ are those in the spaces $\Re^j$, $j = 0, 1/2, 1, 3/2 \ldots$. Thus $SU(2)$ is the "quantum mechanical symmetry group," the symmetry group for the quantum mechanical states.

Implicitly this same result was obtained at the beginning of the Sect 2.3 where representations were obtained for the commutation relations of angular momentum (2.3.1). These commutation relations are the same as the commutation relations of the Lie algebra $SO(3)$ or of the enveloping algebra $\mathscr{E}(SO(3))$ and the enveloping algebra $\mathscr{E}(SU(2))$. For every integer or half-integer value of j, $j = 0, 1/2, 1, 3/2, \ldots$ there is a $(2j + 1)$-dimensional representation space $\Re^j$ in which this Lie algebra acts like the algebra of angular momentum operators.

It indeed happens that for every integer *and* half-integer value of j, there exist quantum physical systems with angular momentum $\hbar j$. Every value of angular momentum allowed by quantum theory is also realized in nature: angular momentum J_i and, specifically, the angular momentum of an extended object about its center of mass or the spin $S_i = J_i - L_i$ in (2.2.15) can have either integer or half-integer values. Electrons, protons, nuclei, and baryons are but a few of many examples of such systems in non-relativistic and relativistic physics. It is remarkable that from the probability interpretation of quantum mechanics (2.6.1) it follows that spin can take integer and half-integer values, and that this is borne out by nature.

Whereas the orbital angular momentum L_i of a mass point can be defined in terms of its momentum P_k and position Q_j, the general angular momentum J_i is an independent observable defined solely by its commutation relation (2.3.1) as linear operators in a linear space. The observables J_i have their origins in the groups of transformations of space-time: they are the "generators" of transformations in the space of quantum physical states that represent the rotations R (of the detection apparatus relative to the preparation apparatus) in three-dimensional space, $x_i \to (Rx)_i$. Similarly, the momenta P_i are the generators of space translations $x_i \to x_i + a_i$, and the Q_i are related to the generators mQ_i of pure Galilean transformations $x_i \to x_i + v_i t$, where m is the mass. The Hamiltonian is the generator of time translations $t \to t + \tau$. These transformations form the Galilean group of symmetry transformations for the non-relativistic space-time. A similar situation holds for the relativistic space-time where the symmetry group is the Poincaré group.

2.7 Summary

The classical orbital angular momentum $\mathbf{l}$ of a mass point with position $\mathbf{x}$ and momentum $\mathbf{p}$ is

$$\mathbf{l} = \mathbf{x} \times \mathbf{p}\,.$$

Using the correspondence principle, the quantum mechanical orbital angular momentum **L** is

$$\mathbf{L} = \mathbf{Q} \times \mathbf{P} ,$$

where **Q** and **P** are, respectively, the position and momentum operators for the mass point. Using the Heisenberg equations of motion, the quantum orbital angular momentum operators are found to obey the algebra

$$[L_i, L_j] = i\hbar \, \epsilon_{ijk} L_k \ .$$

The total angular momentum operators J_i are the sum of the orbital angular momentum operators L_i and spin angular momentum operators S_i,

$$J_i = L_i + S_i \ .$$

The operators J_i and S_i each satisfy the same algebra obeyed by the L_i. Consequences of the algebra are determined using algebraic techniques. Since the operator

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2 \equiv J_k J_k$$

commutes with J_i, vectors are chosen that are eigenvectors of $\mathbf{J}^2$. Vectors are also usually arbitrarily chosen to be eigenvectors of J_3. The orthonormal basis vectors $|j, m\rangle$ satisfy

$$\mathbf{J}^2 |j, m\rangle = \hbar^2 (j + 1) |j, m\rangle \, , \quad J_3 |j, m\rangle = \hbar m |j, m\rangle \, ,$$

where $j = 0, 1/2, 1, 3/2, \ldots$ and $m = j, j - 1, \ldots, -j + 1, -j$. Since $\mathbf{J}^2$ commutes with each of the J_i, there exists no operator that can be formed from the J_i that changes the value of j.

The (non-hermitian) raising operator J_+ and lowering operator J_- are defined by

$$J_\pm \equiv J_1 \pm i J_2 \, ,$$

and are introduced because they satisfy

$$J_3 J_\pm |j, m\rangle = \hbar (m \pm 1) J_\pm |j, m\rangle \, ,$$

revealing that $J_\pm |j, m\rangle$ is either proportional to $|j, m \pm 1\rangle$ or is zero. Including proportionality factors,

$$J_\pm |j, m\rangle = \hbar \sqrt{(j + m)(j \pm m + 1)} \ |j, m \pm 1\rangle \, .$$

The space $\mathfrak{R}^j$ is a $(2j+1)$-dimensional space that is the orthogonal direct sum of the one-dimensional spaces $\mathfrak{R}^j_m$ and is written

$$\mathfrak{R}^j = \sum_{m=-j}^{j} \oplus \mathfrak{R}^j_m ,$$

where each space $\mathfrak{R}^j_m$ is spanned by the vector $|j, m\rangle$.

For a quantum rotator the angular momentum operators J_i and the position operators Q_i representing the internuclear separation satisfy the algebra

$$[Q_i, Q_j] = 0 , \qquad [J_i, J_j] = i\hbar\epsilon_{ijk} J_k , \qquad [J_i, Q_j] = i\hbar\epsilon_{ijk} Q_k .$$

The operators $Q_{\pm 1} \equiv \mp\frac{1}{\sqrt{2}}(Q_1 \pm iQ_2)$ do not commute with J_3 so, as a consequence, change m, while neither Q_3 nor $Q_{\pm 1}$ commute with $\mathbf{J}^2$, so all three operators change j.

A rotator is a physical system that possesses only rotational degrees of freedom. The diatomic molecule is a simple physical system that can, in some approximation, be viewed as such a system. The rotator Hamiltonian for a rigid, stationary, spinning diatomic molecule is

$$H = \frac{\mathbf{J}^2}{2I} ,$$

where I is the moment of inertia of molecule. The energy eigenvalues of the above rotator Hamiltonian are

$$E_j = \frac{\hbar^2 j(j+1)}{2I} .$$

The splitting between rotational energy levels is on the order of 10^{-3} eV. Transitions between rotator states with different values of m or different values of ℓ can be caused by the dipole operator $\mathbf{D} = q\mathbf{Q}$ because it neither commutes with $\mathbf{L}^2$ nor with L_3. At higher values of orbital angular momentum, the rotator is no longer rigid. As a result of elastic potential energy, the Hamiltonian becomes

$$H = \frac{\mathbf{J}^2}{2\mu r_{\mathrm{e}}^2} - \frac{(\mathbf{J}^2)^2}{2k\mu^2 r_{\mathrm{e}}^6} + O\left((\mathbf{J}^2)^3\right) ,$$

where r_{e} is the distance between the two nuclei of the diatomic molecule when $j = 0$.

Problems

For Sect. 2.2

2.1 What operator relation results from taking the adjoint of the operator equation $[L_1, L_2] = i\hbar L_3$?

2.2 Using (2.2.5) and (2.2.6), derive $[L_i, Q_j] = i\epsilon_{ijk} Q_k$.

For Sect. 2.3

2.3 Use the commutation relations of angular momentum (2.3.1) to verify that $[\mathbf{J}^2, J_i] = 0$ where $i = 1, 2, 3 = x, y, z$.

2.4 Verify the equation $J_\pm J_\mp = \mathbf{J}^2 - J_3^2 \pm \hbar J_3$.

2.5 Calculate the matrix element $\langle j, m|J_2|j, m\rangle$.

2.6 Calculate the matrix element $\langle j, m|J_1 J_2 - J_2 J_1)|j, m\rangle$. Is the operator in the matrix element hermitian?

2.7 Calculate the matrix element $(J_-|j, m+1\rangle, J_+|j, m-1\rangle)$ using the following two methods:

(a) Calculate $J_-|j, m+1\rangle$ and $J_+|j, n-1\rangle$ directly and then compute the scalar product.
(b) Use the fact that $J_-^\dagger = J_+$ and write the matrix element as $\langle j, m+1|J_+^2|j, m-1\rangle$. Calculate $J_+^2|j, m-1\rangle$, and then determine the scalar product.

2.8 The baryons Σ^+, Σ^0, and Σ^- are (approximately) eigenstates of isospin. Isospin obeys the algebra of angular momentum but apparently has nothing to do with ordinary angular momentum. Assume Σ^+, Σ^0, and Σ^- have isospin $I(=j) = 1$, and third component of isospin $I_3(=m) = 1, 0,$ and -1, respectively.

(a) Draw a weight diagram labeling the states Σ^+, Σ^0, and Σ^-.
(b) Assume that the mass operator M is given by $M = M_0 1 + M_1 I_3$ when M_0 and M_1 are constants. Then for example,

$$m_{\Sigma^+} = \langle 1, 1|(M_0 1 + M_1 I_3|1, 1\rangle.$$

Using $M_{\Sigma^+} = 1189$ Mev and $M_0 = 1193$ Mev, calculate M_0 and M_1.
(c) What is the predicted mass for the baryon Σ^-?

2.9

(a) Show that the matrices σ_i with matrix elements $2\langle j=\frac{1}{2}, m|J_i|j=\frac{1}{2}, m'\rangle$ are given by

$$\sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_2=\begin{pmatrix}0&-1\\1&0\end{pmatrix},\quad \sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

The σ_i are the Pauli spin matrices.

(b) Show that these matrices satisfy

$$\{\sigma_i,\sigma_j\}\equiv\sigma_i\sigma_j+\sigma_j\sigma_i=2\begin{pmatrix}1&0\\0&1\end{pmatrix}\delta_{ij}.$$

2.10 Construct normalized eigenvectors $|j=\frac{1}{2}, m_1\rangle$,

$$|j=1/2, m_1\rangle=\alpha|j=1/2, m=1/2\rangle+\beta|j=1/2, m=-1/2\rangle,$$

such that the $|j, m_1\rangle$ are simultaneous eigenvectors of $\mathbf{J}^2$ and J_1 satisfying

$$\begin{aligned}\mathbf{J}^2|j,m_1\rangle&=\hbar^2 j(j+1)|j,m_1\rangle,\\ J_1|j,m_1\rangle&=\hbar m_1|j,m_1\rangle.\end{aligned}$$

Calculate the allowed values of m_1.

2.11 Verify that (2.3.61) is correct by expressing L_3 as given in (2.3.54) in terms of the spherical operators a_κ given in (2.3.56).

2.12 In the i^{th} Cartesian dimension let $|n_i\rangle$ be an eigenvector of the harmonic oscillator number operator $N_i=a_i^\dagger a_i$ satisfying $N_i|n_i\rangle=n_i|n_i\rangle$.

(a) Using (1.2.49) and (1.2.50) , respectively, calculate $a_i|n_i\rangle$ and $a_i^\dagger|n_i\rangle$.

(b) Using the expression (2.3.54) for L_3, calculate $L_3|n_1, n_2\rangle$ where $|n_1, n_2\rangle=|n_1\rangle\otimes|n_2\rangle$.

(c) When an eigenvector $|m\rangle$ of L_3 is written as a linear combination of the various eigenvectors $|n_1, n_2\rangle$, explain why the sum n_1+n_2 must be the same for each eigenvector $|n_1, n_2\rangle$ in the linear combination.

(d) Calculate the eigenvalue m of the operator L_3 for the eigenvector $|n_1=0, n_2=0\rangle$.

(e) Let the normalized eigenvector $|m\rangle=\alpha|n_1=1, n_2=0\rangle+|n_1=0, n_2=1\rangle$ where α and β are constants. Calculate the two allowed values of m and, in terms of eigenstates of the two-dimensional harmonic oscillator, determine the normalized eigenstate associated with each value of m.

2.13 Let $|m\rangle$ be an eigenstate of L_3 that satisfies $L_3|m\rangle=\hbar\, m\,|m\rangle$. Express $|m\rangle$ in terms of two-dimensional harmonic oscillator states $|n_1,\ n_2\rangle$ as follows: $|m\rangle=$

$\alpha|2,\ 0\rangle + \beta\,|1,\ 1\rangle + \gamma\,|0,\ 2\rangle$, where α, β, and γ are constants. Calculate the three allowed values of m and, in terms of eigenstates of the two-dimensional harmonic oscillator, determine the normalized eigenstate associated with each value of m.

For Sect. 2.4

2.14 Show that the following four operators all commute: J_3, $\mathbf{J}^2 = J_i J_i$, $\mathbf{Q}^2 = Q_i Q_i$ and $\mathbf{Q}\cdot\mathbf{J} = Q_i J_i$, where the repeated index i is summed from 1 to 3. Explain why states of a rotator can be labeled by the eigenvalues of these four operators and, perhaps, others.

2.15 Using only (2.4.1b) determine the third component m of the angular momentum of the vector $Q_0|r, k_0, j, m\rangle$ where $Q_0 = Q_3$.

2.16 Calculate the commutator $[J_3, \mp\frac{1}{\sqrt{2}}(Q_1 \pm i Q_2)] = [J_3, Q_{\pm 1}]$ with the aid of (2.4.1b). Use the result to determine the third component of angular momentum of the vector $Q_{\pm 1}|r, k_0, j, m\rangle$.

2.17 Just using words without any calculations, explain why $\epsilon_{ikj}\epsilon_{ikm} = 2\delta_{jm}$.

2.18 Calculate $[\mathbf{J}^2, Q_1]$ and $[\mathbf{J}^2, Q_2]$. Verify that the second relation follows from the first by cyclically permuting 1, 2 and 3 (i.e., by making the substitution 1→ 2, 2→ 3, 3→ 1). From your result for $[\mathbf{J}^2, Q_2]$ cyclically permute 1, 2 and 3 to write the result for $[\mathbf{J}^2, Q_3]$.

2.19 Calculate the commutator $[\mathbf{J}^2, \frac{1}{\sqrt{2}}(Q_1 - i Q_2)]$ with the aid of (2.4.1). Use the result to determine the angular momentum of the vector

$$\frac{1}{\sqrt{2}}(Q_1 - i Q_2)|r, k_0, j, m = -j\rangle.$$

2.20 What spaces $\Re^{\ell'}_{m'}$ are obtained when the operator$(Q_3)^2$ acts on the space $\Re^{\ell}_{m}$?

2.21 What spaces $\Re^{\ell'}_{m'}$ are obtained when the operators $(Q_\pm)^2$ act on the space $\Re^{\ell}_{m}$?

2.22 The classical quantity corresponding to $\mathbf{Q}\cdot\mathbf{J}$ is $\mathbf{r}\cdot\mathbf{l}$. Calculate the numerical value of $\mathbf{r}\cdot\mathbf{l}$ for the rigid rotator.

2.23

(a) Using the values of b_{HCl} and d_{HCl} given in (2.4.59), calculate the internuclear distance r_{e} at $\ell = 0$ and the spring constant k for the diatomic molecule HCl.
(b) Calculate the increase in the internuclear distance r as a function of ℓ.

2.24

(a) Consider electromagnetic dipole decay of a diatomic molecule from the energy level $E_{\ell+1}$ to E_ℓ. Show that the frequency $\nu_{\ell,\ell+1}$ of the photon emitted in such a decay is given by

$$\nu_{\ell,\ell+1} = \frac{h}{4\pi^2 I}(\ell + 1).$$

(b) Using the data in the table below, in which the values $\ell+1$ and the corresponding experimental values for frequencies $\nu_{\ell,\ell+1}$ in the far infrared for HCl[17] are given, construct a graph of $\nu_{\ell,\ell+1}$ vs. $\ell + 1$ and determine the moment of inertia I of the HCl molecule from the slope of the graph.

$\ell + 1$	$\nu_{\ell,\ell+1}(\mathrm{cm}^{-1})$
4	83.03
5	104.1
6	124.30
7	145.03
8	165.51
9	185.86
10	206.38
11	226.50

(c) What is the energy of the $\ell = 1$ eigenstate of the HCl molecule?

For Sect. 2.5

2.25 Verify (2.5.3).

2.26

(a) Calculate $\left(\Delta_{|j,m\rangle} J_1\right)\left(\Delta_{|j,m\rangle} J_2\right)$ and then specialize to the case $m = j$.

(b) Make the *incorrect* assumption that $\mathbf{J}^2|j, j\rangle = \hbar^2 j^2 |j, j\rangle$ and show that this implies $\langle j, j|J_1^2|j, j\rangle = \langle j, j|J_2^2|j, j\rangle = 0$. How do these results lead to a contradiction with the Heisenberg uncertainty principle?

---

[17] G. Herzberg, *Molecular Spectra and Molecular Structure* (D. Van Nostrand: New York,1966).

Chapter 3
Combinations of Quantum Physical Systems

3.1 Introduction

The structure of quantum systems is studied by “looking at them” with light or with other quantum systems such as electrons that are usually more fundamental and have less structure than the physical system being investigated. If a quantum particle has no internal structure and is a point object, it is fundamental and is called an elementary particle.

For all other particles, whether they appear to be fundamental or composite depends on the energies used to investigate them. Atoms at scattering energies lower than the first excited energy level appear to be fundamental, structureless objects that can be described as elementary particles. Molecules also appear to be structureless at sufficiently low energies. But as the energy used to probe the quantum system increases, structure becomes evident. For example, the structure of the molecular rotator begins to manifest itself at energies of about 10^{-3} eV (Chap. 2, Sect. 2.4), and the structure of the oscillator appears at energies of 0.1 eV (Chap. 1, Sect. 1.3.2).

As a result of Rutherford’s scattering experiments, an atom or molecule is viewed as consisting of positive nuclei with electrons swirling at an average distance of about 10^{-10} m, a distance that is large compared with the “size” of the nucleus, which is approximately 10^{-15} m. In his experiments Rutherford showed that most alpha particles directed toward atoms were barely deflected while a few were deflected at very large angles. He thus concluded that most of the mass of an atom is concentrated in a small nucleus rather than being uniformly distributed throughout the atom. The nucleus appears to be point-like when examined with a probing projectile that has an energy less than 10^3 eV. As the energy of the probe (photon, alpha particle, electron) used to study the nucleus is increased, the structure of the nucleus manifests itself. At energies of 0.1 MeV $= 10^5$ eV, the structure of a nuclear rotator appears, and at slightly higher energies (1–10 MeV) the nucleus becomes an oscillator. At still higher energies, the individual nucleons (protons and neutrons) in the nucleus reveal themselves. In electron-nucleus scattering, at lower energies the

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9_3

electrons scatter off of the entire nucleus. As the energy of the electrons is increased, the electrons instead scatter off of individual nucleons, which are called hadrons.

Hadrons (neutrons, protons, etc.) appear to be point-like particles at energies below about 1 GeV $= 10^9$ eV, but at higher energies they reveal their structure by becoming relativistic rotators and relativistic oscillators. And at still higher energies they reveal their quark structure. Just as atoms are “made of neutrons, protons and electrons,” hadrons are made of quarks that may or may not be fundamental. In many respects hadrons are similar to molecules because both are understood as being composite. Unlike neutrons, protons, and electrons, which are the constituents of atoms, free quarks, which are the constituents of hadrons, have not been directly detected experimentally at this time in history: they are confined within the hadrons.

The process of deciding which objects in nature are fundamental is ongoing: elementary particles in physics are the exception rather than the rule. At the present time the photon is considered to be an elementary “particle” as are quarks, the electron, muon, tau, and their associated neutrinos. At least experiments involving these particles have not yet revealed any internal structure.

If the energy of an experiment is sufficiently high to reveal that a “particle” has internal structure, physicists build mathematical models in an effort to explain this structure. There are two general approaches: The physicist determines (1) what the system does (vibrates, rotates, etc.) or (2) from what it is made (quarks, protons, neutrons, electrons, nuclei, etc.).

So far in this text, approach (1) has been emphasized. When studying diatomic molecules, their vibrational and rotational spectra were discussed without discussing details regarding the binding of possible constituent electrons and nuclei. In fact, if the diatomic molecules were to be treated strictly from the constituent point of view, it would be necessary to discuss the binding of quarks into neutrons and protons, the binding of neutrons and protons into nuclei, the Coulomb binding of electrons and a nucleus into an atom, and the covalent bonding of atoms to form molecules. Describing molecules in terms of their constituent quarks or in terms of their constituent hadrons and electrons would be very complicated. In the range of molecular physics (10^{-3}–10^{-1} eV), it would be totally unprofitable. Complex molecules are more profitably analyzed in terms of their motions: rotations, vibrations, and the motion of single electrons for the electronic spectra of molecules.

Studying a physical system in terms of constituents (what the system is made from) is often more productive if, in some approximation, the system can be thought of as being made from a small number of objects. This is the approach taken when analyzing the hydrogen atom as a bound state of a proton and an electron. Of course, this description of the hydrogen atom is only approximate: the proton itself is not a fundamental object, and the electron, which may or may not be fundamental, has intrinsic properties such as spin and a magnetic moment. But at the scale of atomic energies and at an accuracy that does not reveal the fine structure and magnetic moments of the constituents, the simple, two-particle model works very well in predicting the energy levels of the hydrogen atom.

In Sect. 3.2 it will be shown that the Hamiltonian of a free particle moving in three dimensions can be thought of as the sum of the Hamiltonians of three free particles, each moving in one dimension. As an extension of this idea, it will then be shown that a Hamiltonian for two interacting particles can be split into a sum of two, one-body Hamiltonians: the Hamiltonian for the motion of the center of mass and the Hamiltonian for motion relative to the center of mass.

In Sect. 3.3 the properties of vectors in three-dimensional space are used to motivate the use of the direct product of linear, scalar-product spaces, the essential mathematical concept required for combining quantum physical systems. The defining mathematical properties of these spaces are then discussed. The examples of Sect. 3.2 are generalized, and Fundamental Postulate IV is formulated for combining two (or more) quantum physical systems. The procedure of dividing two-particle motion into the motion of its constituents and then rewriting the motion in terms of center-of-mass motion and motion relative to the center of mass as discussed in Sect. 3.2 is shown to actually be an example of the use of of Fundamental Postulate IV. Finally, the concept of "complete sets of commuting observables" or "complete sets of commuting operators" is introduced. The mechanism is discussed by which experimental observations both provide guidance and impose constraints on the choice of "complete sets of commuting observables."

In Sect. 3.4 Postulate IV is applied to a specific physical example. When a diatomic molecule is studied with energies $\lesssim 10^{-2}$ eV, only rotational states are excited and the molecule can be described as a rotator. When the energy is increased to about 0.1 eV, the molecule begins vibrating with the result that new energy levels are observed, each of which contains an entire rotational band. The transition of the diatomic molecule from a rotator to a vibrating-rotator provides a specific example of the application of direct-product spaces and the identification of complete sets of commuting operators.

Combining two physical systems with respective angular momenta j_a and j_b is discussed in Sect. 3.6. The combined rotational motion is described by a direct-product space. But angular momentum bases $|j_a, m_a\rangle \otimes |j_b, m_b\rangle$ that are the direct product of the angular momentum bases $|j_a, m_a\rangle$ and $|j_b, m_b\rangle$ of the two constituents, respectively, typically are not physical. A basis is said to be physical if it consists of eigenstates in which the physical system appears. Stationary states (states that do not change in time) are always eigenstates—or mixtures of eigenstates—of the Hamiltonian H, so physical systems such as atoms and molecules are eigenstates of the Hamiltonian. Linear combinations of bases $|j_a, m_a\rangle \otimes |j_b, m_b\rangle$ yield new vectors that are eigenstates of the Hamiltonian, total angular momentum and 3-component of total angular momentum. The transition coefficients used to construct physical bases are called Clebsch-Gordan or Wigner coefficients and are discussed in detail.

3.2 Separation of a Hamiltonian

3.2.1 Introduction

The examples in this section are used to motivate and provide the background to understand the important concept of the direct-product space, a new mathematical tool needed to define the combination of two or more quantum systems. The mathematical notation will first be used here intuitively before being defined in Sect. 3.3.1 below. To introduce the concept of direct-product spaces, it is noted that a Hamiltonian of a non-interacting particle in three-dimensional space can be thought of as the sum of three Hamiltonians, each describing a non-interacting particle moving in a one-dimensional space. As a second example, a two-body Hamiltonian is considered that describes two particles in three-dimensional space interacting via a potential that depends only on the relative positions of the particles. By changing to center-of-mass coordinates and relative coordinates, the two-body Hamiltonian can be written as the sum of two, one-body Hamiltonians, one describing the motion of the center of mass and the other the relative motion resulting from an interaction potential. The technique for separating the two-body Hamiltonian into a sum of two, one-body Hamiltonians is important in its own right: it drastically simplifies the two-body problem by reducing it to two, one-body problems.

3.2.2 A Non-interacting Particle in Three-Dimensional Space Viewed as Three Non-interacting Particles, Each Moving in a One-Dimensional Space

A non-interacting, spinless elementary particle in three-dimensional space is characterized by its mass m and is described by the Hamiltonian

$$H = \frac{1}{2m}\mathbf{P}^2 = \frac{1}{2m}\left[(P_1)^2 + (P_2)^2 + (P_3)^2\right] . \tag{3.2.1}$$

The momentum operators $P_i,\ i = 1, 2, 3$, fulfill the commutation relations

$$\left[P_i,\ P_j\right] = 0 . \tag{3.2.2}$$

The particle in three-dimensional space can be considered as a combination of three particles, each moving in one-dimensional space and described, respectively, by the Hamiltonians

$$H_1 = \frac{P_1^2}{2m} ,\ H_2 = \frac{P_2^2}{2m} ,\ H_3 = \frac{P_3^2}{2m} . \tag{3.2.3}$$

The eigenvectors of the operator P_i are denoted $|p_i\rangle$ and satisfy

$$P_i|p_i\rangle = p_i|p_i\rangle \;, \quad -\infty < p_i < \infty \;, \quad i = 1, 2, 3 \,. \tag{3.2.4}$$

A basis system for the Hamiltonian (3.2.1) is

$$|\mathbf{p}\rangle \;\equiv\; |p_1,\, p_2,\, p_3\rangle \;\equiv\; |p_1\rangle \otimes |p_2\rangle \otimes |p_3\rangle \,. \tag{3.2.5}$$

To indicate, for example, that H_1 only operates on the space spanned by $|p_1\rangle$ it is written in the form $H_1 \otimes 1 \otimes 1$. Here, without being mathematically precise, the direct-product symbol $\otimes$ has been introduced to indicate that the spaces spanned by $|p_1\rangle$, $|p_2\rangle$, and $|p_3\rangle$ are independent and that the operator H_1, for example, only operates on the space spanned by $|p_1\rangle$. This intuitive introduction to direct-product spaces provides a foundation for understanding the mathematical discussion in the following section.

Letting H_1 operate on $|\mathbf{p}\rangle$,

$$\begin{aligned} H_1\,|\mathbf{p}\rangle &= (H_1 \otimes 1 \otimes 1)\,(|p_1\rangle \otimes |p_2\rangle \otimes |p_3\rangle) \;, \\ &= H_1\,|p_1\rangle \otimes 1|p_2\rangle \otimes 1|p_3\rangle \,, \\ &= \frac{P_1^2}{2m}|p_1\rangle \otimes |p_2\rangle \otimes |p_3\rangle \,, \\ &= \frac{p_1^2}{2m}\,|\mathbf{p}\rangle \,. \end{aligned} \tag{3.2.6}$$

The importance of the above equation results not from the equation itself, which is almost trivial, but rather from the fact that it introduces the concept of direct-product spaces. If the space of physical states spanned by $|p_1\rangle$ is denoted $\mathscr{H}_1$ etc., then the space of physical states spanned by (3.2.5) is denoted by

$$\mathscr{H} \equiv \mathscr{H}_1 \otimes \mathscr{H}_2 \otimes \mathscr{H}_3 \,. \tag{3.2.7}$$

The space $\mathscr{H}$ is said to be the direct product of the spaces $\mathscr{H}_1$, $\mathscr{H}_2$, and $\mathscr{H}_3$. Operators A, B, and C that operate, respectively, in spaces $\mathscr{H}_1$, $\mathscr{H}_2$, and $\mathscr{H}_3$ are denoted

$$A \equiv A \otimes 1 \otimes 1 \,, \qquad B \equiv 1 \otimes B \otimes 1 \,, \qquad C \equiv 1 \otimes 1 \otimes C \,. \tag{3.2.8}$$

Using direct-product notation, the Hamiltonian (3.2.1) can be written in the form

$$\begin{aligned} H &= H_1 \otimes 1 \otimes 1 \,+\, 1 \otimes H_2 \otimes 1 \,+\, 1 \otimes 1 \otimes H_3 \\ &= \left(\frac{P_1^2}{2m} \otimes 1 \otimes 1 \,+\, 1 \otimes \frac{P_2^2}{2m} \otimes 1 \,+\, 1 \otimes 1 \otimes \frac{P_3^2}{2m}\right) \,. \end{aligned} \tag{3.2.9}$$

Similarly, the momentum operators are

$$P_1 \equiv P_1 \otimes 1 \otimes 1\,, \quad P_2 \equiv 1 \otimes P_2 \otimes 1\,, \quad P_3 \equiv 1 \otimes 1 \otimes P_3\,. \tag{3.2.10}$$

The space $\mathscr{H}_i$ is spanned by the set of continuous eigenvectors $|p_i\rangle$, which implies that for every vector $\phi_i \in \mathscr{H}_i$, $i = 1, 2, 3$, can be written as

$$\phi_i = \int_{-\infty}^{\infty} \mathrm{d}p_i |p_i\rangle\langle p_i|\phi_i\rangle\,, \quad i = 1, 2, 3\,. \tag{3.2.11}$$

Thus the space $\mathscr{H}_i$ is the set of all ϕ_i:

$$\mathscr{H}_i \equiv \left\{ \phi_i | \phi_i = \int_{-\infty}^{\infty} \mathrm{d}p_i |p_i\rangle\langle p_i|\phi_i\rangle \right\}\,, \quad i = 1, 2, 3\,. \tag{3.2.12}$$

An eigenvector $|x_i\rangle$ of the position operator Q_i satisfies

$$Q_i|x_i\rangle = x_i|x_i\rangle\,, \quad i = 1, 2, 3\,, \tag{3.2.13}$$

and is another basis vector that spans the space $\mathscr{H}_i$ of physical states of a non-interacting particle moving in the i^{th} dimension. In analogy with (3.2.11),

$$|x_i\rangle = \int_{-\infty}^{\infty} \mathrm{d}p_i |p_i\rangle\langle p_i|x_i\rangle\,. \tag{3.2.14}$$

From the commutation relation of the momentum and position operator $[Q_1, P_1] = i\hbar\mathbb{1}$ and the mathematical assumptions about the set of functions $\langle p_i|\phi_i\rangle$ that permit the expansion (3.2.11), it is possible to prove that[1]

$$\langle p_1|x_1\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{-ip_1x_1/\hbar}\,. \tag{3.2.15}$$

Corresponding relations are valid for p_2, x_2 and p_3, x_3 so that

$$\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar}\,, \tag{3.2.16}$$

where x and p are, respectively, x_1, p_1; x_2, p_2 or x_3, p_3.

The basis system $|x_i\rangle$ that spans the space $\mathscr{H}_i$ for one-dimensional motion can immediately be generalized to three dimensions, and the space of physical states $\mathscr{H}$

---

[1] See A. Bohm, *Quantum Mechanics: Foundations and Applications 3rd Ed.* Springer, New York, 2008, eq. (7.50).

in (3.2.7) is spanned by the basis system

$$|\mathbf{x}\rangle \equiv |x_1, x_2, x_3\rangle \equiv |x_1\rangle \otimes |x_2\rangle \otimes |x_3\rangle \,. \tag{3.2.17}$$

With the aid of (3.2.14) and (3.2.16), the above relation becomes

$$\begin{aligned} |\mathbf{x}\rangle = \int_{-\infty}^{\infty} \mathrm{d}p_1 \, |p_1\rangle \frac{1}{\sqrt{2\pi\hbar}} e^{-ip_1x_1/\hbar} \otimes \int \mathrm{d}p_2 \, |p_2\rangle \frac{1}{\sqrt{2\pi\hbar}} e^{-ip_2x_2/\hbar} \\ \otimes \int \mathrm{d}p_3 \, |p_3\rangle \frac{1}{\sqrt{2\pi\hbar}} e^{-ip_3x_3/\hbar} \\ = \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p} \, |p_1\rangle \otimes |p_2\rangle \otimes |p_3\rangle (2\pi\hbar)^{-3/2} \, e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} = \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p} \, |\mathbf{p}\rangle\langle\mathbf{p}|\mathbf{x}\rangle \,. \end{aligned} \tag{3.2.18}$$

The transition coefficients $\langle\mathbf{p}|\mathbf{x}\rangle$, which are the components of the (generalized) basis vectors $|\mathbf{x}\rangle$ along the (generalized) basis vectors $|\mathbf{p}\rangle$, are the plane waves in three-dimensional space,

$$\langle\mathbf{p}|\mathbf{x}\rangle = \langle\mathbf{x}|\mathbf{p}\rangle^* = (2\pi\hbar)^{-3/2} e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} \,. \tag{3.2.19}$$

Plane waves are a mathematical idealization that cannot be attained exactly in the laboratory. According to the *Dirac basis expansion*, a physical state $\psi \in \mathscr{H}$ can be expanded either with respect to the basis system $|\mathbf{p}\rangle$ or $|\mathbf{x}\rangle$. Using the completeness relations (A.4.44)

$$1 = \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p} \, |\mathbf{p}\rangle\langle\mathbf{p}| \,, \quad 1 = \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{x} \, |\mathbf{x}\rangle\langle\mathbf{x}| \,, \tag{3.2.20}$$

ψ can be expanded as follows:

$$\psi = \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p} \, |\mathbf{p}\rangle\langle\mathbf{p}|\psi\rangle \,, \quad \psi = \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{x} \, |\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle \,. \tag{3.2.21}$$

The Schrödinger wave function $\psi(\mathbf{x})$ is obtained by taking the (generalized) scalar product of (3.2.21) with $|\mathbf{x}\rangle$:

$$\psi(\mathbf{x}) \equiv \langle\mathbf{x}|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p} \, \langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}|\psi\rangle = (2\pi\hbar)^{-3/2} \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p} \, \langle\mathbf{p}|\psi\rangle e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \,. \tag{3.2.22}$$

The Schrödinger wave function $\psi(\mathbf{x})$ as given in (3.2.22) is a free-particle wave packet that is a superposition of plane waves.

A non-interacting particle in three-dimensional space can be understood mathematically as a combination of three particles, each moving in a different one-dimensional space. Equivalently, the Hamiltonian of a non-interacting particle can be split into three Hamiltonians, each describing motion in a different one-dimensional space. For two non-interacting particles moving in three-dimensional space, the Hamiltonian can naturally be split into the Hamiltonians of its constituents, which can then be further subdivided into the Hamiltonians of six non-interacting particles, each moving in a one-dimensional space. The presence of an interaction complicates the situation because it is usually more difficult, and sometime impossible, to split a Hamiltonian into independent Hamiltonians.

3.2.3 Two Interacting Particles: Center-of-Mass Motion and Relative Motion

Here a bound-state system is discussed in terms of the constituent picture with attention restricted to cases for which the complex system can be approximated as a bound state of two mass points with respective masses m_1 and m_2. An example of such a system is the hydrogen atom, consisting of an electron and proton bound by the Coulomb potential. A second example is the carbon monoxide molecule CO, which is a bound state of carbon and oxygen resulting from covalent bonding.

The classical energy of this two-particle system is

$$E_{\text{classical}} = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + U(\mathbf{x}_2 - \mathbf{x}_1)\,. \tag{3.2.23}$$

In the above equation $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{x}_1$ and $\mathbf{x}_2$ are the respective momenta and positions of m_1 and m_2. The potential energy $U(\mathbf{x}_2 - \mathbf{x}_1)$ binds the mass points and, as is usually the case, is assumed to depend only on the relative position $\mathbf{x}_2 - \mathbf{x}_1$ of the constituents. The quantum Hamiltonian corresponding to (3.2.23) is given by

$$H = \frac{\mathbf{P}_1^2}{2m_1} + \frac{\mathbf{P}_2^2}{2m_2} + U(\mathbf{Q}_2 - \mathbf{Q}_1)\,. \tag{3.2.24}$$

Guided by (3.2.7), the space of physical states $\mathscr{H}$ is the space

$$\mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2\,, \tag{3.2.25}$$

where $\mathscr{H}_1$ is the space of states for the first mass point m_1 in the three-dimensional space, and $\mathscr{H}_2$ is the space of states for the second mass point m_2. The generalized eigenvectors $|\mathbf{p}_1\rangle$ that satisfy

$$\mathbf{P}_1|\mathbf{p}_1\rangle = \mathbf{p}_1|\mathbf{p}_1\rangle\,, \tag{3.2.26}$$

can be used as a basis system for $\mathscr{H}_1$. Similarly, the generalized eigenvectors $|\mathbf{p}_2\rangle$ span the space $\mathscr{H}_2$. A basis system for $\mathscr{H}$ is then

$$|\mathbf{p}_1, \mathbf{p}_2\rangle \equiv |\mathbf{p}_1\rangle \otimes |\mathbf{p}_2\rangle \,. \tag{3.2.27}$$

According to the continuous basis expansion (3.2.11), the vectors ϕ of the space $\mathscr{H}$ can be expanded as

$$\phi = \iint \mathrm{d}^3\mathbf{p}_1 \, \mathrm{d}^3\mathbf{p}_2 \; |\mathbf{p}_1, \mathbf{p}_2\rangle\langle\mathbf{p}_1, \mathbf{p}_2|\phi\rangle \,, \tag{3.2.28a}$$

where the $\langle\mathbf{p}_1, \mathbf{p}_2|\phi\rangle$ are the momentum-space wave functions that depend on the six components of momentum p_{1i}, p_{2i}. An alternative basis for $\mathscr{H}$ is $|\mathbf{x}_1, \mathbf{x}_2\rangle \equiv |\mathbf{x}_1\rangle \otimes |\mathbf{x}_2\rangle$, which is a simultaneous eigenvector of $\mathbf{Q}_1$ and $\mathbf{Q}_2$. Thus every vector ϕ can also be expanded with respect to $|\mathbf{x}_1, \mathbf{x}_2\rangle$,

$$\phi = \iint \mathrm{d}^3\mathbf{x}_1 \, \mathrm{d}^3\mathbf{x}_2 \; |\mathbf{x}_1, \mathbf{x}_2\,\rangle\langle\mathbf{x}_1, \mathbf{x}_2|\phi\rangle \,, \tag{3.2.28b}$$

The continuous expansion coefficients $\langle\mathbf{x}_1, \mathbf{x}_2|\phi\rangle$ are the position-space wave functions; they are functions of the six coordinates x_{1i}, x_{2i}. The set of six coordinates $\{x_{1i}, x_{2i}\}$ span a space that is called the configuration space. The functions $\langle\mathbf{x}_1, \mathbf{x}_2|\phi\rangle$ are also often called the configuration-space wave functions to emphasize that the wave function $\phi(\mathbf{x}_1, \mathbf{x}_2) \equiv \langle\mathbf{x}_1, \mathbf{x}_2|\phi\rangle$ of a two-particle quantum system is not a wave in the three-dimensional coordinate space $\Re_3$ but instead represents a "state" in the six-dimensional "configuration space".

The relationship between the eigenvectors $|\mathbf{x}_1, \mathbf{x}_2\rangle$ and $|\mathbf{p}_1, \mathbf{p}_2\rangle$ is found by replacing ϕ by $|\mathbf{x}_1, \mathbf{x}_2\rangle$ in (3.2.28a),

$$\begin{aligned} |\mathbf{x}_1, \mathbf{x}_2\rangle &= \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p}_1 \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p}_2 \; |\mathbf{p}_1, \mathbf{p}_2\rangle\langle\mathbf{p}_1, \mathbf{p}_2|\mathbf{x}_1, \mathbf{x}_2\rangle \,, \\ &= (2\pi\hbar)^{-3} \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p}_1 \int_{-\infty}^{\infty} \mathrm{d}^3\mathbf{p}_2 \; |\mathbf{p}_1, \mathbf{p}_2\rangle e^{-i(\mathbf{p}_1\cdot\mathbf{x}_1 + \mathbf{p}_2\cdot\mathbf{x}_2)/\hbar} \,. \end{aligned} \tag{3.2.29}$$

The above equation is the generalization of (3.2.19) to the two-particle case. Using direct-product notation, in the direct-product space $\mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2$, the position and momentum operators in (3.2.24) are

$$\mathbf{Q}_1 = \mathbf{Q}_1 \otimes 1 \,, \qquad \mathbf{Q}_2 = 1 \otimes \mathbf{Q}_2 \,, \tag{3.2.30a}$$

$$\mathbf{P}_1 = \mathbf{P}_1 \otimes 1 \,, \qquad \mathbf{P}_2 = 1 \otimes \mathbf{P}_2 \,. \tag{3.2.30b}$$

The operator $U(\mathbf{Q}_2 - \mathbf{Q}_1)$ acts in both spaces $\mathscr{H}_1$ and $\mathscr{H}_2$. As a consequence, it typically *cannot*[2] be written in the form $U = U_1 \otimes 1 + 1 \otimes U_2$. The Hamiltonian, therefore, also cannot be written as a sum of operators that operate either in $\mathscr{H}_1$ or $\mathscr{H}_2$; this will always be the case for an interacting, two-particle system.

It is difficult to determine the eigenvalues of the Hamiltonian (3.2.24) in its present form because it involves six momentum operators P_{1x}, P_{1y}, P_{1z}, P_{2x}, P_{2y}, P_{2z} and six position operators Q_{1x}, Q_{1y}, Q_{1z}, Q_{2x}, Q_{2y}, Q_{2z} characteristic of a two-body problem. The two-particle Hamiltonian can be simplified, however, by writing it in terms of new operators that are chosen in such a way that the two-body Hamiltonian is a sum of two, one-body Hamiltonians. To accomplish this task, first note that the potential is not an arbitrary function of $\mathbf{Q}_1$ and $\mathbf{Q}_2$, but, according to (3.2.24), only depends on the difference $\mathbf{Q}_2 - \mathbf{Q}_1$. As a consequence, the potential depends only on the relative position operator $\mathbf{Q}$ defined by

$$\mathbf{Q} = \mathbf{Q}_2 - \mathbf{Q}_1 \,. \tag{3.2.31}$$

The correspondence between classical and quantum mechanics suggests using the center-of-mass position operator $\mathbf{Q}_{\mathrm{CM}}$, where

$$\mathbf{Q}_{\mathrm{CM}} = \frac{m_1\mathbf{Q}_1 + m_2\mathbf{Q}_2}{M} \,, \quad M = m_1 + m_2 \,. \tag{3.2.32}$$

The momentum operator $\mathbf{P}$ conjugate to the relative position operator $\mathbf{Q}$ in (3.2.31) is

$$\mathbf{P} = \frac{m_1\mathbf{P}_2 - m_2\mathbf{P}_1}{M} \,. \tag{3.2.33}$$

The center-of-mass momentum operator $\mathbf{P}_{\mathrm{CM}}$ conjugate to the center-of-mass position operator $\mathbf{Q}_{\mathrm{CM}}$ in (3.2.32) is

$$\mathbf{P}_{\mathrm{CM}} = \mathbf{P}_1 + \mathbf{P}_2 \,. \tag{3.2.34}$$

As can readily be checked using the three-dimensional generalization of the Heisenberg commutation relations (1.2.11), namely,

$$[Q_i, P_j] = i\hbar\delta_{ij}\mathbb{1}, \quad [P_i, P_j] = 0, \quad [Q_i, Q_j] = 0, \tag{3.2.35}$$

it is straightforward to show that $\mathbf{Q}$ in (3.2.31) and $\mathbf{P}$ in (3.2.33) satisfy $[Q_i, P_j] = i\hbar\delta_{ij}\mathbb{1}$ and that $\mathbf{Q}_{\mathrm{CM}}$ in (3.2.32) and $\mathbf{P}_{\mathrm{CM}}$ in (3.2.34) satisfy $[Q_{\mathrm{CM}i}, P_{\mathrm{CM}j}] = i\hbar\delta_{ij}\mathbb{1}$ as conjugate positions and momenta must.

---

[2] For example, when $U(\mathbf{Q}_2 - \mathbf{Q}_1) = k[(\mathbf{Q}_2 - \mathbf{Q}_1)^2] = k[(\mathbf{Q}_2)^2 + (\mathbf{Q}_1)^2 + 2\mathbf{Q}_1\mathbf{Q}_2]$, the final term prevents U from being written as the direct product $U = U_1 \otimes 1 + 1 \otimes U_2$.

For two operators $A_1 = A_1 \otimes 1$ and $B_2 = 1 \otimes B_2$ it always follows that $[A_1, B_2] = 0$ so that the position and momentum operators for mass point 1 commute with the position and momentum operators for mass point 2. Also, $\mathbf{P}_1$ and $\mathbf{Q}_1$ obey the Heisenberg commutation relations (3.2.35) as do $\mathbf{P}_2$ and $\mathbf{Q}_2$. From the definitions (3.2.31)–(3.2.34), the following commutation relations can be immediately derived. (Problem 3.1):

$$\begin{aligned}
&\left[Q_i, P_j\right] = i\hbar\delta_{ij}\mathbb{1} && \left[Q_{\mathrm{CM}i}, P_{\mathrm{CM}j}\right] = i\hbar\delta_{ij}\mathbb{1} \\
&\left[P_i, P_j\right] = \left[Q_i, Q_j\right] = 0 && \left[P_{\mathrm{CM}i}, P_{\mathrm{CM}j}\right] = \left[Q_{\mathrm{CM}i}, Q_{\mathrm{CM}j}\right] = 0 \\
&\left[P_i, P_{\mathrm{CM}j}\right] = \left[P_i, Q_{\mathrm{CM}j}\right] = 0 && \left[Q_i, P_{\mathrm{CM}j}\right] = \left[Q_i, Q_{\mathrm{CM}j}\right] = 0
\end{aligned} \tag{3.2.36}$$

Equations (3.2.36) reveal that the relative momentum and position operators $\mathbf{P}$ and $\mathbf{Q}$ obey the Heisenberg commutation relations, that the center-of-mass momentum and position operators $\mathbf{P}_{\mathrm{CM}}$ and $\mathbf{Q}_{\mathrm{CM}}$ obey the Heisenberg commutation relations, and that the relative momentum and position operators commute with the center-of-mass momentum and position operators.

The Hamiltonian (3.2.24) can be rewritten in terms of the operators defined in (3.2.31)–(3.2.34) and takes the form (Problem 3.2),

$$H = H_{\mathrm{CM}} + H_{\mathrm{relative}}\,, \tag{3.2.37a}$$

where

$$H_{\mathrm{CM}} = \frac{\mathbf{P}_{\mathrm{CM}}^2}{2M}\,, \quad M = m_1 + m_2\,, \tag{3.2.37b}$$

$$H_{\mathrm{relative}} = \frac{\mathbf{P}^2}{2\mu} + U(\mathbf{Q})\,, \tag{3.2.37c}$$

and the reduced mass μ is given by

$$\mu = (m_1 m_2)/(m_1 + m_2)\,. \tag{3.2.37d}$$

Because $\mathbf{P}_{\mathrm{CM}}$ and $\mathbf{Q}_{\mathrm{CM}}$ both commute with $\mathbf{P}$ and $\mathbf{Q}$ *and* because $\mathbf{P}_{\mathrm{CM}}$ and $\mathbf{Q}_{\mathrm{CM}}$ obey the Heisenberg commutation relations as do $\mathbf{P}$ and $\mathbf{Q}$, the first term in (3.2.37a) is the Hamiltonian of a non-interacting particle with mass M, and the second term is the Hamiltonian of a particle with mass μ in an external potential U.

The space in which H_{CM}, $\mathbf{P}_{\mathrm{CM}}$ and $\mathbf{Q}_{\mathrm{CM}}$ act is the space spanned by the eigenkets $|\mathbf{p}_{\mathrm{CM}}\rangle$ of $\mathbf{P}_{\mathrm{CM}}$ that satisfy

$$\mathbf{P}_{\mathrm{CM}}|\mathbf{p}_{\mathrm{CM}}\rangle = \mathbf{p}_{\mathrm{CM}}|\mathbf{p}_{\mathrm{CM}}\rangle\,, \tag{3.2.38}$$

where $\mathbf{P}_{\mathrm{CM}}$ is defined in (3.2.34). This space is denoted $\mathscr{H}_{\mathrm{CM}}$. Similarly, the space in which H_{relative}, $\mathbf{P}$ and $\mathbf{Q}$ act is the space spanned by $|\mathbf{p}\rangle$ that satisfies

$$\mathbf{P}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle\,, \tag{3.2.39}$$

where $\mathbf{P}$ is defined in (3.2.33). This space is denoted $\mathscr{H}_{\text{relative}}$. The space $\mathscr{H}$ of (3.2.25) can also be written as the direct product

$$\mathscr{H} = \mathscr{H}_{\mathrm{CM}} \otimes \mathscr{H}_{\text{relative}}\,. \tag{3.2.40}$$

Instead of the basis system (3.2.27) or (3.2.29), the following basis system is used:

$$|\mathbf{p}_{\mathrm{CM}}, \mathbf{p}\rangle \equiv |\mathbf{p}_{\mathrm{CM}}\rangle \otimes |\mathbf{p}\rangle\,. \tag{3.2.41}$$

When written in terms of the center-of-mass and relative operators, the Hamiltonian is the sum of H_{CM} and H_{relative}, which act, respectively, in the spaces $\mathscr{H}_{\mathrm{CM}}$ and $\mathscr{H}_{\text{relative}}$. Writing the Hamiltonian in direct product notation,

$$H = H_{\mathrm{CM}} \otimes 1 + 1 \otimes H_{\text{relative}} = \frac{\mathbf{P}_{\mathrm{CM}}^2}{2M} \otimes 1 + 1 \otimes \left[\frac{\mathbf{P}^2}{2\mu} + U(\mathbf{Q})\right]. \tag{3.2.42}$$

By rewriting the Hamiltonian in terms of the center-of-mass momentum operator $\mathbf{P}_{\mathrm{CM}}$ and relative momentum and position operators $\mathbf{P}$ and $\mathbf{Q}$, the two-body Hamiltonian (3.2.24) has been converted into a sum of two, independent, one-body Hamiltonians that are much easier to analyze.

In both (3.2.25) and (3.2.42), the space of physical states $\mathscr{H}$ is written as a direct product of two spaces. In the first case the spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ are, respectively, the space of states for the mass points m_1 and m_2. In the second case the spaces $\mathscr{H}_{\mathrm{CM}}$ and and $\mathscr{H}_{\text{relative}}$ are the space of states for the center-of-mass and relative motion, respectively. While all four spaces $\mathscr{H}_1$, $\mathscr{H}_2$, $\mathscr{H}_{\mathrm{CM}}$, and $\mathscr{H}_{\text{relative}}$ are three dimensional, they are very different spaces.

The basis vectors (3.2.38) are eigenvectors of H_{CM}, but the eigenvectors (3.2.39), which are eigenvectors of $\mathbf{P}$, are not eigenvectors of H_{relative}:

$$[P_i, H_{\text{relative}}] \neq 0 \qquad \text{because} \qquad [P_i, U(\mathbf{Q})] \neq 0 \qquad i = 1, 2, 3\,. \tag{3.2.43}$$

Since a vector can be an eigenvector of two operators iff the two operators commute, $|\mathbf{p}\rangle$ cannot be an eigenvector of H_{relative}. One of the standard problems of quantum mechanics is to determine the eigenvalues of the Hamiltonian H, and rewriting Hamiltonian as the direct product of two Hamiltonians, as is done in (3.2.40), is the first step in solving the problem. Applying H to (3.2.41),

$$H|\mathbf{p}_{\mathrm{CM}}, \mathbf{p}\rangle = \left(\frac{\mathbf{p}_{\mathrm{CM}}^2}{2M}|\mathbf{p}_{\mathrm{CM}}\rangle\right) \otimes |\mathbf{p}\rangle + |\mathbf{p}_{\mathrm{CM}}\rangle \otimes H_{\text{relative}}|\mathbf{p}\rangle\,, \tag{3.2.44}$$

where $\mathbf{p}_{\mathrm{CM}}^2/2M$ is the kinetic energy of the center of mass. Because of (3.2.43), $|\mathbf{p}\rangle$ is not an eigenstate of H_{relative}: when H_{relative} acts on $|\mathbf{p}\rangle$, it transforms $|\mathbf{p}\rangle$ into a continuous linear combination of momentum kets $|\mathbf{p}'\rangle$ with all possible values for p_i'. Thus the use of the basis system $|\mathbf{p}\rangle$ in the space $\mathscr{H}_{\text{relative}}$ is not a practical way to determine the eigenvalues of H_{relative}. To obtain the eigenvalues E of H_{relative}, a basis system of eigenvectors of H_{relative} are required.

Example 3.2.1 Determine the energy eigenvectors and eigenvalues of the Hamiltonian

$$H = \frac{P_1^2}{2m_1} + \frac{P_2^2}{2m_2} + \frac{1}{2}k(Q_2 - Q_1)^2$$

that describes two particles, each of which is moving in a one-dimensional space and interacting via a harmonic oscillator potential with a spring constant k.

Solution Rewriting the Hamiltonian in terms of the relative position operator (3.2.31), the relative momentum operator (3.2.33), and the center-of-mass momentum operator (3.2.34), from (3.2.37) the Hamiltonian takes the form

$$H = \frac{P_{\mathrm{CM}}^2}{2M} + \frac{P^2}{2\mu} + \frac{1}{2}kQ^2 = \frac{P_{\mathrm{CM}}^2}{2M} + H_{\text{oscillator}} = \frac{P_{\mathrm{CM}}^2}{2M} \otimes 1 + 1 \otimes H_{\text{oscillator}} .$$

A suitable basis system is $|p_{\mathrm{CM}}\rangle \otimes |E_n\rangle$ where the $|E_n\rangle$ are the eigenvectors (1.2.36) of a harmonic oscillator with mass μ and spring constant k. The eigenvalues of H are immediately calculated as follows:

$$\begin{aligned} H|p_{\mathrm{CM}}\rangle \otimes |E_n\rangle &= \left(\frac{P_{\mathrm{CM}}^2}{2M}|p_{\mathrm{CM}}\rangle\right) \otimes |E_n\rangle + |p_{\mathrm{CM}}\rangle \otimes \left(\frac{P^2}{2\mu} + \frac{1}{2}kQ^2\right)|E_n\rangle \\ &= \left(\frac{p_{\mathrm{CM}}^2}{2M} + E_n\right)|p_{\mathrm{CM}}\rangle \otimes |E_n\rangle \end{aligned}$$

The energy eigenvalues E_n are given in (1.2.40), $E_n = \hbar\omega(n + 1/2)$ where $\omega = (k/\mu)^{1/2}$ and $n = 0, 1, 2, \ldots$.

Eigenvectors of H_{relative}, temporarily denoted $|E \cdots\rangle$, satisfy

$$H_{\text{relative}}|E \cdots\rangle = E|E \cdots\rangle . \tag{3.2.45}$$

In Example 3.2.1, because the Hamiltonian $H_{\text{oscillator}} \equiv H_{\text{relative}}$ acts in one spatial dimension, the eigenvalue equation for the operator H_{relative} completely determines the energy eigenvectors $|E_n\rangle$ except for a phase factor, $|E_n\rangle \rightarrow e^{i\phi(E_n)}|E_n\rangle$, and a normalization factor, $|E_n\rangle \rightarrow \text{constant}\,|E_n\rangle$. When H_{relative} acts in more spatial dimensions, the eigenvalue equation for the operator H_{relative} does not fully determine the vectors $|E \cdots\rangle$. In order to calculate the eigenvectors of H_{relative}

(up to phase and normalization factors), the $|E\cdots\rangle$ of (3.2.45) are also required to be eigenvectors of a set of additional operators $\{A_k | k = 1, 2, \ldots\}$ satisfying $[H_{\text{relative}}, A_k] = 0$. The best choice for the set of operators $\{A_k\}$ depends on the specific problem.

As is often the case, the potential energy operator $U(\mathbf{Q})$ commutes with the angular momentum operator $L_i = \epsilon_{ijk} P_j Q_k$,

$$[U(\mathbf{Q}), L_i] = 0\,. \tag{3.2.46}$$

Because $[P^2, L_i] = 0$, it then follows that

$$[H_{\text{relative}}, L_i] = 0\,. \tag{3.2.47}$$

From the above equation and (2.3.6), respectively, $\mathbf{L}^2$ and L_3 commute with H_{relative} and with each other so a suitable choice for the additional operators $\{A_k\}$ is $\{\mathbf{L}^2, L_3\}$, where L_3 is the component of $\mathbf{L}$ in any desired direction.

A set of commuting, selfadjoint operators whose eigenvalues completely label the basis vectors is called *a complete set of commuting operators* or *a complete set of commuting observables.* When studying a quantum physical system, the two most important questions are as follows: (1) What is a complete set of commuting operators for the system? (2) What are the algebraic relations among the various operators?

For a quantum physical system with a classical analogue, the number of operators in a complete set of commuting observables equals the number of degrees of freedom in the corresponding classical system. In Example 3.2.1 on the previous page, each particle moves in one spatial dimension, so there are two degrees of freedom. Thus the two operators, $P_{\text{CM}}^2/2M$ and $H_{\text{oscillator}}$ constitute a complete set of commuting operators. For the quantum mechanical Kepler problem describing a spinless electron in a Coulomb field, the electron moves in three spatial dimensions, so there are three degrees of freedom. The operators $\mathscr{H}_{\text{relative}}$, $\mathbf{L}^2$, and L_3 form a complete set of commuting operators. Taking electron spin into consideration, as will be discussed later, the spin-angular momentum has to be added to the orbital angular momentum to form total angular momentum operators. The additional spin degree of freedom requires that the complete set of commuting operators be enlarged.

To determine the eigenvalues of H_{relative} for the quantum mechanical Kepler problem, it is convenient to choose the eigenvectors $|E\cdots\rangle = |E, \ell, m\rangle$ as the basis vectors for the space $\mathscr{H}_{\text{relative}}$ that fulfill the eigenvalue equations

$$\left[\frac{\mathbf{P}^2}{2\mu} + U(\mathbf{Q})\right] |E, \ell, m\rangle = E|E, \ell, m\rangle\,, \tag{3.2.48a}$$

$$\mathbf{L}^2|E, \ell, m\rangle = \hbar^2 \ell(\ell+1)|E, \ell, m\rangle\,, \tag{3.2.48b}$$

$$L_3|E, \ell, m\rangle = \hbar m|E, \ell, m\rangle\,. \tag{3.2.48c}$$

From the discussion of angular momentum, in particular from (2.3.33) and (2.3.34), ℓ can take integer values, $\ell = 0, 1, 2, \cdots$, and $m = -\ell, -\ell+1, \ldots, \ell$. In general the eigenvalues E, ℓ and m are all interrelated just as ℓ and m are. The possible eigenvalues of E in (3.2.48a) obviously depend on H_{relative} and also depend on the boundary conditions.

Boundary conditions are a requirement on the set of acceptable solutions. For example, only those $|E, \ell, m\rangle$ for which the values of E are a discrete set $\{E_n | n = 0, 1, 2, \cdots\}$ might be deemed acceptable. A standard method for solving the eigenvalue equation (3.2.48a) is to take the "scalar product" of $|E, \ell, m\rangle$ with the eigenkets $|\mathbf{x}\rangle$ of the operator $\mathbf{Q}$ of (3.2.31), $\mathbf{Q}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle$:

$$\langle\mathbf{x}| \left(\frac{P_i P_i}{2\mu} + U(\mathbf{Q}) \right) |E, \ell, m\rangle = E \langle\mathbf{x}|E, \ell, m\rangle \,. \tag{3.2.49}$$

Equation (3.2.49) is then converted into a differential equation that can be solved by expanding $|E, \ell, m\rangle$ in terms of momentum eigenstates $|\mathbf{p}\rangle$, just as ψ was expanded in terms such states in (3.2.21):

$$\langle\mathbf{x}|P_i|E, \ell, m\rangle = \int_{-\infty}^{\infty} \mathrm{d}^3 p \langle\mathbf{x}|P_i|\mathbf{p}\rangle\langle\mathbf{p}|E, \ell, m\rangle = \int_{-\infty}^{\infty} \mathrm{d}^3 p \; p_i \langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}|E, \ell, m\rangle \,. \tag{3.2.50}$$

Using (3.2.19) the above equation can be rewritten in the form

$$\langle\mathbf{x}|P_i|E, \ell, m\rangle = (2\pi\hbar)^{-\frac{3}{2}} \int_{-\infty}^{\infty} \mathrm{d}^3 p \; p_i e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \langle\mathbf{p}|E, \ell, m\rangle \,. \tag{3.2.51}$$

Since

$$\frac{\hbar}{i}\frac{\partial}{\partial x_i} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} = \frac{\hbar}{i}\frac{i p_i}{\hbar} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} = p_i e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \,,$$

(3.2.51) becomes

$$\begin{aligned}\langle\mathbf{x}|P_i|E, \ell, m\rangle &= (2\pi\hbar)^{-\frac{3}{2}} \int_{-\infty}^{\infty} \mathrm{d}^3 p \frac{\hbar}{i}\frac{\partial}{\partial x_i} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \langle\mathbf{p}|E, \ell, m\rangle \\ &= \frac{\hbar}{i}\frac{\partial}{\partial x_i} \int_{-\infty}^{\infty} \mathrm{d}^3 p \langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}|E, \ell, m\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x_i} \langle\mathbf{x}|E, \ell, m\rangle \,. \end{aligned} \tag{3.2.52}$$

Using the same logic that led to (3.2.52), it immediately follows that for any integer n,

$$\langle\mathbf{x}|(P_i)^n|E, \ell, m\rangle = \left(\frac{\hbar}{i}\frac{\partial}{\partial x_i} \right)^n \langle\mathbf{x}|E, \ell, m\rangle \,. \tag{3.2.53}$$

With the aid of (3.2.53), the energy eigenvalue equation (3.2.49) becomes

$$\left[\frac{1}{2\mu}\left(\frac{\hbar}{i}\frac{\partial}{\partial x^i}\right)^2 + U(\mathbf{x})\right]\langle\mathbf{x}|E,\ell,m\rangle = E\langle\mathbf{x}|E,\ell,m\rangle\,. \tag{3.2.54}$$

The Schrödinger wave functions $\psi_{n,\ell,m}(\mathbf{x}) = \langle\mathbf{x}|E_n,\ell,m\rangle$ are required to obey specific boundary conditions just as are solutions to any differential equation. For example, the solutions could be required to be Lebesgue square integrable functions—or Hilbert space functions—of the variables $\mathbf{x}$. In most physical applications, solutions are at least chosen to be smooth, rapidly decreasing functions. With such boundary conditions (3.2.54) often has solutions for a discrete set of values of E that are negative, $E_1, E_2, E_3, \ldots$ and a continuous set of values E that are non-negative, $0 \le E < \infty$.

The set of all possible energy eigenvalues of (3.2.54) is called the energy spectrum or the spectrum of the energy operator H. Note that this set is determined both by the differential operator in (3.2.54) and by the boundary conditions. A typical example of this spectrum is the spectrum of the hydrogen atom or the quantum mechanical Kepler problem.

Example 3.2.2 Determine the lowest energy eigenvalue and corresponding Schrödinger wave function for a hydrogen atom in a two-dimensional space described by the Hamiltonian

$$H = \frac{P_x^2 + P_y^2}{2\mu} - \frac{e^2}{4\pi\epsilon_0\sqrt{Q_x^2 + Q_y^2}}\,.$$

The hydrogen atom has a reduced mass μ, the magnitude of the electron and proton charges are e, and the interaction results from a Coulomb potential in two spatial dimensions.

Solution Since the Hamiltonian is spherically symmetric (Problems 3.3 and 3.4),

$$[H, L_3] = 0\,,$$

it is possible to choose eigenvectors $|E,m\rangle$ where E and m are, respectively, the eigenvalues of H and L_3. Using the same procedure that led to (3.2.54), the Schrödinger equation in two dimensions is

$$\left[-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - \frac{e^2}{4\pi\epsilon_0\sqrt{x^2+y^2}})\right]\langle x,y|E,m\rangle = E\langle x,y|E,m\rangle\,,$$

and the eigenvalue equation $L_3|E, m\rangle = \hbar m|E, m\rangle$ similarly becomes

$$\left(x\frac{\hbar}{i}\frac{\partial}{\partial y} - y\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\langle x, y|E, m\rangle = m\hbar\langle x, y|E, m\rangle\,.$$

Rewriting the above two equations in terms of polar coordinates, $x = r\cos\phi$ and $y = r\sin\phi$ (Problems 3.7 and 3.8),

$$\left\{-\frac{\hbar^2}{2\mu}\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2}\right] - \frac{e^2}{4\pi\epsilon_0 r}\right\}\langle r, \phi|E, m\rangle = E\langle r, \phi|E, m\rangle\,,$$

$$\frac{\hbar}{i}\frac{\partial}{\partial\phi}\langle r, \phi|E, m\rangle = \hbar m\langle r, \phi|E, m\rangle\,.$$

The second of the above two equations is the analog of (3.2.48c) and makes it possible to calculate $\frac{\partial^2}{\partial\phi^2}\langle r, \phi|E, m\rangle$ in the energy eigenvalue equation. Writing $\langle r, \phi|E, m\rangle = R(r)\Phi(\phi)$, the solution to the second equation is

$$\Phi(\phi) = e^{im\phi}\,,$$

and the energy eigenvalue equation becomes

$$\left\{-\frac{\hbar^2}{2\mu}\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{m^2}{r^2}\right] - \frac{e^2}{4\pi\epsilon_0 r}\right\}R(r) = ER(r)\,.$$

One boundary condition requires that the solution be invariant under rotations of 2π, $\Phi(\phi + 2\pi) = \Phi(\phi)$, implying that m is an integer.

Since the term in the energy eigenvalue equation proportional to m^2 that multiplies $R(r)$ is positive, the term increases the energy. The smallest energy thus occurs when m is zero. The solution corresponding to this energy is of the form $R(r) = e^{-r/a_0}$, where a_0 is a positive constant. Note that the solution is square-integrable. Using this trial solution, the eigenvalue equation for the energy becomes

$$\left(-\frac{\hbar^2}{2\mu\, a_0^2} - E\right) + \frac{1}{r}\left(\frac{\hbar^2}{2\mu a_0} - \frac{e^2}{4\pi\epsilon_0}\right) = 0\,.$$

The above equation must be satisfied for all values of r, so the two terms in parentheses must vanish independently. Consequently,

$$a_0 = \frac{2\pi\epsilon_0\hbar^2}{\mu e^2}\,, \qquad E = -\frac{2\mu e^4}{(4\pi\epsilon_0\hbar)^2}\,.$$

The (unnormalized) solution is

$$\langle r, \phi|E, m\rangle = R(r)\Phi(\phi) = e^{-r/a_0}\,e^{im\phi}\,.$$

Example 3.2.3 Determine the form of the Schrödinger equation (3.2.54) for the spherically invariant potential $U(r)$ ($r = \sqrt{x^2 + y^2 + z^2}$) in the three-dimensional space, using spherical coordinates: $(x, y, z) = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta)$. Express the final result in terms of the operator $\mathbf{L}^2$ (or its eigenvalues) and derivatives with with respect to a variable r.

Solution Taking the square of the angular momentum operator $\mathbf{L} = \mathbf{Q} \times \mathbf{P}$ we obtain with the help antisymmetric symbol (2.2.4)

$$\begin{aligned}\mathbf{L}^2 &= (\mathbf{Q} \times \mathbf{P})^2 = \epsilon_{ijk} Q_j P_k \epsilon_{ilm} Q_l P_m = (\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) Q_j P_k Q_l P_m \\ &= Q_j P_k Q_j P_k - Q_j P_k Q_k P_j = Q_j (Q_j P_k - i\hbar\delta_{jk}) P_k - Q_j P_k (P_j Q_k + i\hbar\delta_{jk}) \\ &= Q_j Q_j P_k P_k - 2i\hbar Q_j P_j - (Q_j P_j Q_k P_k - 3i\hbar) = \mathbf{Q}^2\mathbf{P}^2 + i\hbar \mathbf{Q}\cdot\mathbf{P} - (\mathbf{Q}\cdot\mathbf{P})^2.\end{aligned}$$

Now let us calculate $\mathbf{Q} \cdot \mathbf{P}$ and $(\mathbf{Q} \cdot \mathbf{P})^2$

$$\frac{\partial}{\partial r} = \frac{x}{r}\frac{\partial}{\partial x} + \frac{y}{r}\frac{\partial}{\partial y} + \frac{z}{r}\frac{\partial}{\partial z} \Rightarrow \mathbf{Q}\cdot\mathbf{P} = r\frac{\hbar}{i}\frac{\partial}{\partial r} \text{ and } (\mathbf{Q}\cdot\mathbf{P})^2 = -\hbar^2 r^2 \frac{\partial^2}{\partial r^2} - \hbar^2 r \frac{\partial}{\partial r},$$

so we obtain

$$\mathbf{P}^2 = \frac{1}{r^2}((\mathbf{Q}\cdot\mathbf{P})^2 - i\hbar(\mathbf{Q}\cdot\mathbf{P}) + \mathbf{L}^2)$$

and the Schrödinger equation (3.2.54) in the spherical coordinates takes the form

$$\left(-\frac{\hbar^2}{2\mu}\Big(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{\ell(\ell+1)}{r^2}\Big) + U(r)\right)\langle r, \theta, \phi | n, \ell, m\rangle = E\langle r, \theta, \phi | n, \ell, m\rangle \tag{3.2.55}$$

In three-dimensional space solutions of the Schrödinger equation (3.2.55) are of the form $\langle r, \theta, \phi | n, \ell, m\rangle \equiv R_{n,\ell}(r) Y_\ell^m(\theta, \phi)$. For the hydrogen atom the first few radial functions $R_{n,\ell}(r)$ and spherical harmonics $Y_\ell^m(\theta, \phi)$ are tabulated below ($r_B = 4\pi\epsilon_0\hbar^2/\mu e^2$ is the Bohr radius and $L_{n-\ell-1}^{2l+1}(\rho)$ is a generalized Laguerre polynomial of degree $n - \ell - 1$) (Tables 3.1 and 3.2):

The solutions have been normalized so that the probability of finding the electron somewhere is unity. The formula for all energy levels for a hydrogen atom are similarly found to be

$$E_n = -\frac{e^2}{8\pi\epsilon_0 r_B}\frac{1}{n^2} = -\frac{\mu e^4}{2(4\pi\epsilon_0\hbar)^2}\frac{1}{n^2} = -\frac{13.6\,\text{eV}}{n^2}, \text{ and } -\infty < E_n < 0. \tag{3.2.56}$$

Table 3.1 Normalized, radial wave functions for the hydrogen atom

n	ℓ	$R_{n,\ell}(r)$
1	0	$\frac{2}{(r_B)^{3/2}}\,\mathrm{e}^{-r/r_B}$
2	0	$\frac{1}{(2r_B)^{3/2}}\left(2-\frac{r}{r_B}\right)\mathrm{e}^{-r/2r_B}$
2	1	$\frac{1}{\sqrt{3}(2r_B)^{3/2}}\frac{r}{r_B}\,\mathrm{e}^{-r/2r_B}$
3	0	$\frac{2}{3(3r_B)^{3/2}}\left[3-2\frac{r}{r_B}+\frac{2}{9}\left(\frac{r}{r_B}\right)^2\right]\mathrm{e}^{-r/3r_B}$
3	1	$\frac{4\sqrt{2}}{9(3r_B)^{3/2}}\frac{r}{r_B}\left(1-\frac{1}{6}\frac{r}{r_B}\right)\mathrm{e}^{-r/3r_B}$
3	2	$\frac{2\sqrt{2}}{27\sqrt{5}(3r_B)^{3/2}}\left(\frac{r}{r_B}\right)^2\mathrm{e}^{-r/3r_B}$
...		
n	ℓ	$\left(\frac{2}{nr_B}\right)^{3/2}\sqrt{\frac{(n-l-1)!}{2n(n+l)!}}\mathrm{e}^{-r/(nr_B)}L^{2\ell+1}_{n-\ell-1}(\frac{2r}{nr_B})$

Table 3.2 Spherical harmonics

ℓ	m	$Y_\ell^m(\theta,\phi)$
0	0	$\frac{1}{2\sqrt{\pi}}$
1	0	$\frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta$
1	± 1	$\mp\frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta\, e^{\pm i\phi}$

Because the calculation in Example 3.2.2 was performed in two spatial dimensions instead of three, there the solution for E differs from the expression for the ground-state energy (3.2.56) of the hydrogen atom by a factor of 4. The energy values E_n for the hydrogen atom are plotted in Fig. 3.1 on the following page.

Although more complicated, the situation is similar for other atoms and for the electronic spectra of molecules. The ideal harmonic oscillator, for which $H_{\text{relative}} = \mathbf{P}^2/(2\mu) + (k/2)\mathbf{Q}^2$, has only a discrete spectrum $E_n = \hbar\omega(n + 1/2)$. But as discussed in Chaps. 1 and 2, for realistic oscillators such as vibrating diatomic molecules, the harmonic oscillator potential $U(\mathbf{x}) = k\mathbf{x}^2/2$ is only an approximation that is good for low energies.

In general, as is the case for the Alkali atoms, the discrete eigenvalues of H depend on the angular momentum quantum number ℓ, $E = E_{n,\ell}$, and the possible values of ℓ depend on the value of n. For a given n, the orbital angular momentum can take the values $\ell = 0, 1, \ldots, n-1$ and the energy values for a given n are $E_{n,\ell=0}, E_{n,\ell=1}, \ldots, E_{n,\ell=n-1}$. In contrast, for the Coulomb potential the discrete energy eigenvalues as given in (3.2.56) are independent of ℓ. But when the effects of the spin of the electron are included, the energy levels split into sublevels that depend on ℓ. When a Hamiltonian H_{relative} satisfies (3.2.47), it is said to be rotationally invariant or possess rotational symmetry, and the energy values will not depend on the quantum number m.

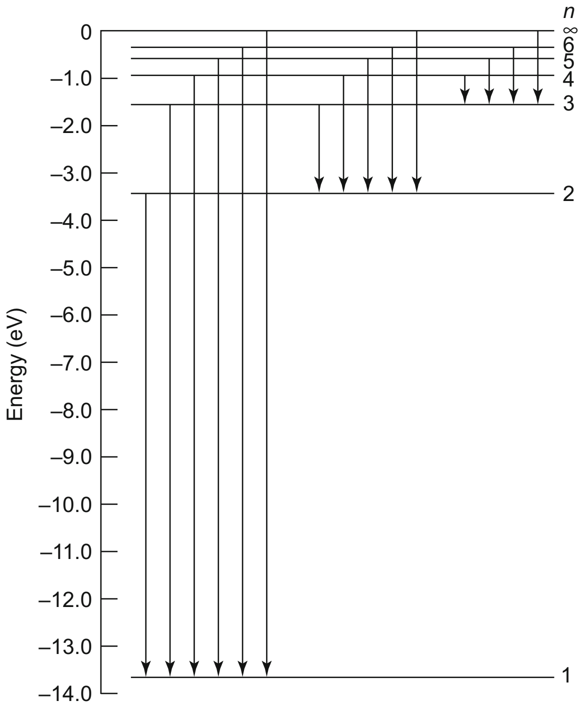


Fig. 3.1 The energy spectrum of the Hamiltonian H_{relative} for the hydrogen atom as given in (3.2.56)

The eigenvectors $|E, \ell, m\rangle$ of (3.2.48) form a basis system for the space of states of the relative motion. This means that every state vector $\phi \in \mathscr{H}_{\text{relative}}$ can be written as

$$\phi = \sum_{n,\ell,m} |E_{n\ell}, \ell, m\rangle\langle E_{n\ell}, \ell, m|\phi\rangle + \int_0^\infty \mathrm{d}E \sum_{\ell,m} |E, \ell, m\rangle\langle E, \ell, m|\phi\rangle , \tag{3.2.57}$$

where the sum over ℓ extends over $\ell = 0, 1, 2, \ldots,$ the sum over m extends over $m = -\ell, -\ell + 1, \ldots, \ell$. Thus the sum over E extends over all discrete values $E_{n,\ell}$. Similarly, the integral over E extends over all values E in the continuous range $0 \leq E < \infty$.

Example 3.2.4 Determine the energy eigenvectors and eigenvalues for the Hamiltonian

$$H = \frac{P_1^2}{2m} + \frac{P_2^2}{2m} + \frac{1}{2}kQ_1^2 + \frac{1}{2}kQ_2^2 + \frac{1}{2}k_1 (Q_2 - Q_1)^2 ,$$

that describes two particles, each moving in one-dimensional space. Here the particles, which could, for example, have different charges, are not assumed to be identical.

Solution Since the potential does not depend solely on the relative coordinate $Q_2 - Q_1$ earlier results of this section cannot be used. The Hamiltonian is separated by introducing new position operators Q_α and Q_β that are linear combinations of the original position operators,

$$Q_\alpha \equiv \alpha\,(Q_2 - Q_1)\ , \qquad Q_\beta \equiv \beta\,(Q_2 + Q_1)\ ,$$

where α and β are constants. An easy method for finding the conjugate momentum operators P_α and P_β is to use the position representation and the chain rule for differentiation. In the position representation,

$$x_\alpha = \alpha(x_2 - x_1)\,, \qquad x_\beta = \beta(x_2 + x_1)\,,$$

and the conjugate momentum operators are

$$P_\alpha = -i\hbar\frac{\partial}{\partial x_\alpha}\ , \quad P_\beta = -i\hbar\frac{\partial}{\partial x_\beta}\ .$$

Using the chain rule for differentiation,

$$-i\hbar\frac{\partial}{\partial x_1} = -i\hbar\frac{\partial x_\alpha}{\partial x_1}\frac{\partial}{\partial x_\alpha} - i\hbar\frac{\partial x_\beta}{\partial x_1}\frac{\partial}{\partial x_\beta} = -\alpha\left(-i\hbar\frac{\partial}{\partial x_\alpha}\right) + \beta\left(-i\hbar\frac{\partial}{\partial x_\beta}\right)\ .$$

The above equation implies

$$P_1 = -\alpha P_\alpha + \beta P_\beta\ .$$

A similar calculation yields

$$P_2 = \alpha P_\alpha + \beta P_\beta\ .$$

Rewriting H in terms of Q_α, Q_β, P_α, and P_β,

$$H = \frac{2\alpha^2}{2m}P_\alpha^2 + \frac{2\beta^2}{2m}P_\beta^2 + \frac{1}{2}\left(\frac{k}{2\alpha^2} + \frac{k_1}{\alpha^2}\right)Q_\alpha^2 + \frac{1}{2}\frac{k}{2\beta^2}Q_\beta^2\ .$$

Choosing $\alpha^2 = \beta^2 = \frac{1}{2}$, the Hamiltonian can be written in the form

$$H = H_\alpha + H_\beta$$

where

$$H_\alpha = \frac{P_\alpha^2}{2m} + \frac{1}{2}(k+2k_1)Q_\alpha^2 \quad \text{and} \quad H_\beta = \frac{P_\beta^2}{2m} + \frac{1}{2}kQ_\beta^2 .$$

Since the Hamiltonian is the sum of two independent harmonic oscillator Hamiltonians, each in a one-dimensional space, the energy eigenvectors are $|E_i^\alpha\rangle \otimes |E_j^\beta\rangle$. The $|E_n^\alpha\rangle$ are the eigenvectors (1.2.36) of a harmonic oscillator with mass m and spring constant $k+2k_1$ and the $|E_j^\beta\rangle$ are eigenvectors for a harmonic oscillator with mass m and spring constant k. The energy eigenvalues are calculated as follows:

$$H|E_i^\alpha\rangle \otimes |E_j^\beta\rangle = \left[\frac{P_\alpha^2}{2m} + \frac{1}{2}(k+2k_1)Q_\alpha^2\right]|E_i^\alpha\rangle \otimes |E_j^\beta\rangle$$

$$+|E_i^\alpha\rangle\otimes\left[\frac{P_\beta^2}{2m} + \frac{1}{2}kQ_\beta^2\right]|E_j^\beta\rangle = \left[\hbar\omega_\alpha\left(i+\frac{1}{2}\right) + \hbar\omega_\beta\left(j+\frac{1}{2}\right)\right]|E_i^\alpha\rangle\otimes|E_j^\beta\rangle .$$

In the above formula $\omega_\alpha = [(k+2k_1)/m]^{1/2}$, $\omega_\beta = (k/m)^{1/2}$, and $i, j = 0, 1, 2, 3, \ldots$.

For classical systems the decomposition of a physical system into constituents is understood to be totally different from the decomposition of the motions (degrees of freedom). In quantum mechanics these decompositions are very similar. For the decomposition into constituents (particles), $\mathscr{H}$ is the direct product of constituent spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ as given in (3.2.25). For the decomposition into center-of-mass motion and relative motion, $\mathscr{H}$ is the direct product of $\mathscr{H}_{\text{CM}}$ and $\mathscr{H}_{\text{relative}}$ as given in (3.2.37a).

3.3 Combination of Quantum Physical Systems

3.3.1 *Mathematical Preliminaries*

The general prescription for the combination of two or more quantum physical systems can be conjectured from the results of the previous section. The mathematical tool needed to give a general description of two quantum physical systems is the new mathematical concept of the *direct product* or *tensor product* of two linear spaces. The idea of direct-product spaces is introduced by considering two vectors (also called rank one tensors) **a** and **b** in the usual three-dimensional space $\Re_3$:

$$\mathbf{a} = \sum_{i=1}^{3} a_i \mathbf{e}_i , \quad \text{and} \quad \mathbf{b} = \sum_{i=1}^{3} b_i \mathbf{e}_i , \tag{3.3.1}$$

where the $\mathbf{e}_i$ are a set of three orthonormal vectors. Writing the two vectors next to each other, a rank-two tensor $T \equiv \mathbf{ab}$ is obtained:

$$\mathbf{ab} = \sum_{i=1}^{3} a_i \mathbf{e}_i \sum_{j=1}^{3} b_j \mathbf{e}_j = \sum_{i=1}^{3} \sum_{j=1}^{3} a_i b_j \mathbf{e}_i \mathbf{e}_j \, , \tag{3.3.2}$$

To stress the fact that the vectors are just written next to each other, mathematicians write $\mathbf{ab} \equiv \mathbf{a} \otimes \mathbf{b}$ and $\mathbf{e}_i \mathbf{e}_j \equiv \mathbf{e}_i \otimes \mathbf{e}_j$. Then, in this new notation (3.3.2) becomes

$$\mathbf{a} \otimes \mathbf{b} = \sum_{i=1}^{3} \sum_{j=1}^{3} a_i b_j \mathbf{e}_i \otimes \mathbf{e}_j \, , \tag{3.3.3}$$

and $\mathbf{e}_i \otimes \mathbf{e}_j$ is called the direct product of the vectors $\mathbf{e}_i$ and $\mathbf{e}_j$. The new space obtained from the direct product is denoted $\Re_3 \otimes \Re_3$ and is spanned by $\mathbf{e}_i \otimes \mathbf{e}_j$, which is a (real) linear space. The following example establishes that $\Re_3 \otimes \Re_3$ possesses the property (A.2.1b), which is one of the properties required of a (real) linear space.

Example 3.3.1 Let T and $\tilde{T}$ be two tensors in the space $\Re_3 \otimes \Re_3$,

$$T = \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij} \, \mathbf{e}_i \otimes \mathbf{e}_j \, , \quad \tilde{T} = \sum_{i=1}^{3} \sum_{j=1}^{3} \tilde{T}_{ij} \, \mathbf{e}_i \otimes \mathbf{e}_j \, .$$

Show that addition is associative. That is, show that $T + \tilde{T} = \tilde{T} + T$.

Solution

$$T + \tilde{T} = \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij} \, \mathbf{e}_i \otimes \mathbf{e}_j + \sum_{i=1}^{3} \sum_{j=1}^{3} \tilde{T}_{ij} \, \mathbf{e} \otimes \mathbf{e}_j$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{3} \tilde{T}_{ij} \, \mathbf{e}_i \otimes \mathbf{e}_j + \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij} \, \mathbf{e}_i \otimes \mathbf{e}_j = \sum_{i=1}^{3} \sum_{j=1}^{3} (\tilde{T}_{ij} + T_{ij}) \, \mathbf{e}_i \otimes \mathbf{e}_j = \tilde{T} + T$$

In a similar manner, it is possible to show that the remaining defining relations (A.2.2b)–(A.2.10b) for a (real) linear space are satisfied. Because the components of vectors in three-dimensional space are real, the constants "a" and "b" that appear in (A.2.1b)–(A.2.9b) must be real. Thus the space being discussed here is a real linear space, not a general linear space.

Instead of taking the direct product of $\Re_3$ with itself, the direct product can be taken of two different, finite-dimensional, linear, scalar-product spaces Φ_M and Ψ_N. The M-dimensional space Φ_M is spanned by ϕ_μ, $\mu = 1, 2, \ldots, M$, and the N-dimensional space Ψ_N is spanned by ψ_ν, $\nu = 1, 2, ..., N$. An element T in the direct-product space $\Phi_M \otimes \Psi_N$ is written in the form

$$T = \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu}\, \phi_\mu \otimes \psi_\nu \,. \tag{3.3.4}$$

Since ϕ_μ is a basis in Φ_M and ψ_ν is a basis in Ψ_N, $\phi_\mu \otimes \psi_\nu$ is a basis in the space $\Phi_M \otimes \Psi_N$ with dimension $M \cdot N$.

A scalar product can be defined on the direct-product space $\Phi_M \otimes \Psi_N$. If $(\phi_\mu, \phi_{\mu'})_\Phi$ denotes the scalar product in Φ_M and $(\psi_\nu, \psi_{\nu'})_\Psi$ denotes the scalar product in Ψ_N, the scalar product of two tensors T and $\tilde{T}$ in the space $\Phi_M \otimes \Psi_N$ is defined by

$$\begin{aligned}(T, \tilde{T}) &= \left(\sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu} \phi_\mu \otimes \psi_\nu \,, \sum_{\mu'=1}^{M} \sum_{\nu'=1}^{N} \tilde{T}_{\mu'\nu'}\, \phi_{\mu'} \otimes \psi_{\nu'} \right) \\ &= \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} \sum_{\mu'=1}^{M} \sum_{\nu'=1}^{N} T^*_{\mu\nu}\, \tilde{T}_{\mu'\nu'} (\phi_\mu, \phi_{\mu'})_\Phi (\psi_\nu, \psi_{\nu'})_\Psi \,. \end{aligned} \tag{3.3.5}$$

It is straightforward to show that the scalar product defined in the above equation satisfies all the properties (A.2.13c)–(A.2.16c) required of a scalar product in a linear, scalar-product space (Problem 3.10).

Example 3.3.2 Verify that (3.3.5) satisfies (A.2.13c). That is, verify that

$$(T, T) \ge 0 \,.$$

Solution Making the replacement $\tilde{T} \to T$ in (3.3.5) ,

$$(T, T) = \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} \sum_{\mu'=1}^{M} \sum_{\nu'=1}^{N} T^*_{\mu\nu}\, T_{\mu'\nu'} (\phi_\mu, \phi_{\mu'})_\Phi (\psi_\nu, \psi_{\nu'})_\Psi \,.$$

Without loss of generality, the basis vectors can be chosen to be orthonormal. Making this choice,

$$(T, T) = \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} \sum_{\mu'=1}^{M} \sum_{\nu'=1}^{N} T^*_{\mu\nu}\, T_{\mu'\nu'}\, \delta_{\mu,\mu'}\, \delta_{\nu,\nu'} = \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T^*_{\mu\nu}\, T_{\mu\nu} \ge 0 \,.$$

If A_Φ is a linear operator on the space Φ_M and B_Ψ is a linear operator on the space Ψ_N, the corresponding linear operators in $\Phi_M \otimes \Psi_N$ are defined to be

$$A \equiv A_\Phi \otimes 1 \,, \qquad B \equiv 1 \otimes B_\Psi \,. \tag{3.3.6}$$

The action of A and B on the tensor T are defined to be

$$AT = (A_\Phi \otimes 1) \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu}\, \phi_\mu \otimes \psi_\nu \equiv \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu} (A_\Phi \phi_\mu) \otimes \psi_\nu \,,$$
$$BT \equiv \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu}\, \phi_\mu \otimes (B_\Psi \psi_\nu)\,. \tag{3.3.7}$$

The action of the operator $C \equiv A_\Phi \otimes B_\Psi$ on T is defined by

$$CT = (A_\Phi \otimes B_\Psi) \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu}\, \phi_\mu \otimes \psi_\nu \equiv \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu} (A_\Phi \phi_\mu) \otimes (B_\Psi \psi_\nu)\,. \tag{3.3.8}$$

It is not difficult to show that the operator C defined above is a linear operator in the direct-product space $\Phi_M \otimes \Psi_N$. That is, the operator C possesses the properties (A.3.1b)–(A.3.5b) required of linear operators (Problem 3.11).

Example 3.3.3 Let T and $\tilde{T}$ be tensors in $\Phi_M \otimes \Psi_N$ and let the action of C on T be defined by (3.3.8). Show that $C(T + \tilde{T}) = CT + C\tilde{T}$.

Solution

$$C(T + \tilde{T}) = (A_\Phi \otimes B_\Psi) \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} (T_{\mu\nu} + \tilde{T}_{\mu\nu}) \phi_\mu \otimes \psi_\nu$$
$$= \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} (T_{\mu\nu} + \tilde{T}_{\mu\nu}) (A_\Phi \phi_\mu) \otimes (B_\Psi \psi_\nu)$$
$$= \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} T_{\mu\nu} (A_\Phi \phi_\mu) \otimes (B_\Psi \psi_\nu) + \sum_{\mu=1}^{M} \sum_{\nu=1}^{N} \tilde{T}_{\mu\nu} (A_\Phi \phi_\mu) \otimes (B_\Psi \psi_\nu) = CT + C\tilde{T}\,.$$

If A_Φ and A'_Φ are linear operators on the space Φ_M while B_Ψ and B'_Ψ are linear operators in the space Ψ_N, then (Problem 3.12)

$$A_\Phi A'_\Phi \otimes B_\Psi B'_\Psi = (A_\Phi \otimes B_\Psi)(A'_\Phi \otimes B'_\Psi)\,. \tag{3.3.9}$$

In the direct-product space every linear operator C is a linear combination of direct products of operators,

$$C = \sum a_i A^i_\Phi \otimes B^i_\Psi \,, \quad a_i \in \mathbb{C}\,. \tag{3.3.10}$$

The direct product of infinite-dimensional, linear, scalar-product spaces is a generalization of the finite-dimensional case obtained by taking the limit $M, N \to \infty$.

3.3.2 Basic Postulate for the Combination of Quantum Systems

The mathematical properties of direct-product spaces have been set forth for $M \times N$ dimensional spaces. To describe the combination of two quantum physical systems, direct-product spaces will be generalized to infinite (and even continuously-infinite) dimensions.

For the case of two interacting mass points considered in the previous section, let $\mathscr{H}_1 = \Phi$ be the space spanned by the (generalized) momentum basis vectors $|\mathbf{p}_1\rangle$ and let $\mathscr{H}_2 = \Psi$ be the space spanned by the (generalized) momentum basis vectors $|\mathbf{p}_2\rangle$, where the (generalized) vectors $|\mathbf{p}_1\rangle$ and $|\mathbf{p}_2\rangle$ describe the first and second mass points, respectively. The space of physical states $\mathscr{H}$ is the direct product of $\mathscr{H}_1$ and $\mathscr{H}_2$, $\mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2$, spanned by the basis system

$$|\mathbf{p}_1, \mathbf{p}_2\rangle \equiv |\mathbf{p}_1\rangle \otimes |\mathbf{p}_2\rangle\,, \tag{3.3.11}$$

and is the space of physical states of the two, mass-point system. Operators for the combined system are operators in the direct-product space $\mathscr{H}$. The momentum operators for mass points 1 and 2 are, respectively, $P_1 \otimes 1$ and $1 \otimes P_2$.

As discussed in Sect. 3.2, (3.3.11) is not the only basis system for the space $\mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2$. Another basis system is given by the vectors $|\mathbf{x}_1, \mathbf{x}_2\rangle \equiv |\mathbf{x}_1\rangle \otimes |\mathbf{x}_2\rangle$. In order to perform calculations, in Sect. 3.2.3 it was shown to be advantageous to consider the two-particle system in terms of the center-of-mass motion and the relative motion about the center of mass. That is, the combined motion is described by the direct product of the center-of-mass motion described by $\mathscr{H}_{\text{CM}}$ and the relative motion described by $\mathscr{H}_{\text{relative}}$. Thus The basis system of $\mathscr{H}$ adapted to the combination of $\mathscr{H}_{\text{CM}}$ and $\mathscr{H}_{\text{relative}}$ is

$$|\mathbf{p}_{\text{CM}}, \mathbf{p}\rangle = |\mathbf{p}_{\text{CM}}\rangle \otimes |\mathbf{p}\rangle\,, \tag{3.3.12}$$

where $|\mathbf{p}_{\text{CM}}\rangle$ is a basis of $\mathscr{H}_{\text{CM}}$ and $|\mathbf{p}\rangle$ is a basis of $\mathscr{H}_{\text{relative}}$. The basic postulate for combining quantum physical systems is an immediate generalization of the case just discussed:

3.3.3 Combination of Quantum Systems: Direct-Product Spaces—Fundamental Postulate IV

Let one physical system be described by an algebra of operators $\{A_\Phi\}$ in the space Φ, and a second physical system be described by an algebra of operators $\{B_\Psi\}$ in the space Ψ. The direct-product space $\Phi \otimes \Psi$ is the space of physical states of the combination of the two systems. The observables for the combined system are operators in the direct-product space and are of the form (3.3.10). Observables in the first system alone are given by $A_\Phi \otimes 1$, and observables in the second system alone are given by $1 \otimes B_\Psi$, where 1 is the identity operator.

3.3.4 Complete Sets of Commuting Observables

A set of commuting, hermitian operators with eigenvalues that completely specify a (generalized) basis for a system is, following Dirac, called *a complete set of commuting operators* or *a complete set of commuting observables* (c.s.c.o.) for the system.

The eigenvalues of a c.s.c.o. are called quantum numbers that can take a discrete set of values, a continuous set of values, or both. For example, for the two point-particles of Sect. 3.2, the c.s.c.o. are the six momentum operators, either of the individual particles OR of the center-of-mass and the relative motion:

$$P_{1i},\ P_{2i} \quad \text{or} \quad P_{\mathrm{CM}_i},\ P_i, \quad \text{where } i = 1, 2, 3\,. \tag{3.3.13}$$

Their respective eigenvectors are

$$|\mathbf{p}_1\,,\ \mathbf{p}_2\rangle \quad \text{or} \quad |\mathbf{p}_{\mathrm{CM}}\,,\ \mathbf{p}\rangle\,. \tag{3.3.14}$$

Alternatively, instead of the six momentum operators, the six position operators could be chosen

$$Q_{1i}\,,\ Q_{2i} \text{ or } Q_{\mathrm{CM}_i}\,,\ Q_i\,, \tag{3.3.15}$$

with respective eigenvectors

$$|\mathbf{x}_1,\ \mathbf{x}_2\rangle \quad \text{or} \quad |\mathbf{x}_{\mathrm{CM}},\ \mathbf{x}\rangle\,. \tag{3.3.16}$$

As already stated in (3.2.28a) and (3.2.28b), each vector ϕ in the direct-product space $\phi \subseteq \mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2 = \mathscr{H}_{\mathrm{CM}} \otimes \mathscr{H}_{\mathrm{relative}}$ can be expanded in terms of $|\mathbf{p}_1,\ \mathbf{p}_2\rangle$ or $|\mathbf{x}_1,\ \mathbf{x}_2\rangle$. Expansion in terms of direct-product basis functions follows immediately from (3.2.21). Each basis vector $|\mathbf{x}_1,\ \mathbf{x}_2\rangle$ can be expanded with respect to the system of basis vectors $|\mathbf{p}_1,\ \mathbf{p}_2\rangle$ as in (3.2.29). Conversely, each basis vector $|\mathbf{p}_1,\ \mathbf{p}_2\rangle$ can be expanded with respect to the system of basis vectors $|\mathbf{x}_1,\ \mathbf{x}_2\rangle$,

$$|\mathbf{p}_1, \mathbf{p}_2\rangle = \iint \mathrm{d}^3x_1\ \mathrm{d}^3x_2 |\mathbf{x}_1, \mathbf{x}_2\rangle\langle\mathbf{x}_1, \mathbf{x}_2|\mathbf{p}_1, \mathbf{p}_2\rangle\,, \tag{3.3.17}$$

where the transition coefficients between these two basis systems are given by

$$\langle\mathbf{x}_1, \mathbf{x}_2|\mathbf{p}_1, \mathbf{p}_2\rangle = (2\pi\hbar)^{-3/2} e^{i(\mathbf{x}_1\cdot\mathbf{p}_1)/\hbar} (2\pi\hbar)^{-3/2} e^{i(\mathbf{x}_2\cdot\mathbf{p}_2)/\hbar}\,. \tag{3.3.18}$$

The above transition coefficients are the complex conjugates of those in (3.2.29), $\langle \mathbf{x}_1, \mathbf{x}_2 | \mathbf{p}_1, \mathbf{p}_2 \rangle = \langle \mathbf{p}_1, \mathbf{p}_2 | \mathbf{x}_1, \mathbf{x}_2 \rangle^*$. Similarly, the basis vectors for center-of-mass and relative position satisfy

$$\langle \mathbf{x}_{\mathrm{CM}}, \mathbf{x} | \mathbf{p}_{\mathrm{CM}}, \mathbf{p} \rangle = (2\pi\hbar)^{-3/2} e^{i\mathbf{x}_{\mathrm{CM}} \cdot \mathbf{p}_{\mathrm{CM}}/\hbar} (2\pi\hbar)^{-3/2} e^{i\mathbf{x}\cdot\mathbf{p}/\hbar} . \tag{3.3.19}$$

The general state vector ϕ can be expanded in terms of any one of the four sets of basis vectors (3.3.14) and (3.3.16). For example, choosing the center-of-mass position basis $|\mathbf{x}_{\mathrm{CM}}, \mathbf{x}\rangle$,

$$\phi = \iint \mathrm{d}^3 x_{\mathrm{CM}} \, \mathrm{d}^3 x \; |\mathbf{x}_{\mathrm{CM}}, \mathbf{x}\rangle \langle \mathbf{x}_{\mathrm{CM}}, \mathbf{x} | \phi \rangle . \tag{3.3.20a}$$

Similarly, choosing the center of mass momentum basis system $|\mathbf{p}_{\mathrm{CM}}, \mathbf{p}\rangle$,

$$\phi = \iint \mathrm{d}^3 p_{\mathrm{CM}} \, \mathrm{d}^3 p \; |\mathbf{p}_{\mathrm{CM}}, \mathbf{p}\rangle \langle \mathbf{p}_{\mathrm{CM}}, \mathbf{p} | \phi \rangle . \tag{3.3.20b}$$

Thus for quantum systems consisting of two point particles that may or may not interact, four different sets of continuous quantum numbers have already been found: $\{\mathbf{p}_1, \mathbf{p}_2\}$, $\{\mathbf{x}_1, \mathbf{x}_2\}$, $\{\mathbf{p}_{\mathrm{CM}}, \mathbf{p}\}$ and $\{\mathbf{x}_{\mathrm{CM}}, \mathbf{x}\}$. The set of possible values that each quantum number can take is called the spectrum of the quantum number. Consequently, the spectrum of the operator P_{1i} is the continuous spectrum $-\infty \leq p_{1i} \leq +\infty$ with the spectra of the other operators described similarly. Each of the four sets consists of six quantum numbers regardless of which set of operators (3.3.13) or (3.3.15) is chosen. Since six numbers are required to characterize two classical mass points, from the correspondence principle it follows that six operators constitute a c.s.c.o. for the two quantum mass points. In general, the number of operators in the c.s.c.o. equals the number of degrees of freedom in the classical system.

For the one-dimensional oscillator of Chap. 1, Sect. 1.2, the energy quantum number n with the spectrum $n = 0, 1, 2, 3 \ldots$ was used to label the basis system. The basis vectors labeled by this quantum number are the energy eigenvectors $|E_n\rangle = |n\rangle$, and the c.s.c.o. consists of the single operator N or, equivalently, the energy operator $H = \hbar\omega(N + 1/2)$. Instead of the eigenvectors $|n\rangle$, the eigenvectors $|x\rangle$ of the position operator satisfying $Q|x\rangle = x|x\rangle$ with the continuous spectrum $-\infty \leq x \leq +\infty$ could have been used. Equivalently, the momentum basis vectors $|p\rangle$ could have been used.

3.3.5 The Rotating Dumbbell

For the rigid rotator discussed in Chap. 2, Sect. 2.2 the two operators $\mathbf{J}^2$ and J_3 constitute a c.s.c.o. The eigenvectors of the rotator are $|j, j_3\rangle$. To complete the set of operators describing the rigid rotator, the operators J_1 and J_2 must be adjoined to the operators $\mathbf{J}^2$ and J_3, where $\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2$. The operators J_i obey the algebra of angular momentum.

As discussed in Chap. 2, Sect. 2.4, in the description of the rigid, rotating dumbbell, which is a combination of two structureless mass points connected by a rigid, massless rod of length r, the angular momentum operators J_i were used, but were insufficient. Additionally, the vector operators Q_i were introduced that are the components "pointing" along the intranuclear axis and satisfying the condition $\mathbf{Q}^2 = Q_i Q_i = r^2$, which is a constant for a rigid, rotating dumbbell. In addition to the intrinsic or relative position operator $\mathbf{Q}$ of the two mass points given by (3.2.31), there is the center-of-mass position operator $\mathbf{Q}_{\mathrm{CM}}$ given by (3.2.32) that was ignored in Chap. 2, Sect. 2.4.

If the dumbbell possesses only orbital angular momentum of the two mass points relative to the center of mass, then $J_i = L_i = \epsilon_{ijk} Q_j P_k$. Since Q_i and Q_j commute, $J_i \cdot Q_i = \epsilon_{ijk} Q_i Q_j P_k = 0$. For a general, axially symmetric rotator, corresponding to a dumbbell with spinning mass points, there is also intrinsic angular momentum (spin) k_0 around the dumbbell axis. Then $J_i \cdot Q_i = \pm k_0 r$, where $k_0 = 0, 1/2, 1, 3/2 \ldots$ and the plus (minus) sign is chosen if the spin is parallel (antiparallel) to $\mathbf{r}$. The c.s.c.o. consists of the set $\{J_3,\ \mathbf{J}^2 = J_i J_i,\ J_i \cdot Q_i,\ Q_i Q_i\}$, and in the center-of mass frame the basis vectors $|j_3,\ j,\ k_0,\ r\rangle$ are labeled by four quantum numbers.

At low energies on the order of 10^{-3} eV, the dumbbell axis is rigid, implying that $r =$ constant. But at higher energies, which could occur in collisions with electrons that have energies on the order of 10^{-1} eV, the dumbbell axis starts vibrating. Then $\mathbf{Q}^2 = Q_i Q_i = r^2$ is no longer constant. In addition to $\mathbf{J}^2$ and J_3, $Q = \sqrt{\mathbf{Q}^2}$ becomes an operator. The rigid dumbbell molecule has become a vibrating-rotating or a rotating-vibrating molecule. For the simplest case where the angular momentum $k_0 = 0$ around the intranuclear axis, the dumbbell molecule rotates about its center of mass and performs harmonic oscillations along the intranuclear axis.

Whether or not the rotations of the vibrating molecule are observed depends on the resolution of the spectrometer used to detect the emitted or absorbed spectral lines. For coarse resolution corresponding to energies $> 10^{-2}$ eV, only one broad spectral line is observed with peaks at the following frequency:

$$\text{Frequency}: \nu = \frac{\Delta E}{2\pi\hbar} = 6.42 \times 10^{13}\,\mathrm{s}^{-1} \text{ where } \Delta E = |E_n - E_{n\pm 1}| = 0.265\,\mathrm{eV} \tag{3.3.21a}$$

Expressing the frequency as a wave number by dividing by c.[3]

$$\text{Wave number}: \nu = \frac{\Delta E}{2\pi\hbar\, c} = 2140\,\mathrm{cm}^{-1} \tag{3.3.21b}$$

---

[3] In molecular physics wave number is expressed in cm^{-1} and is denoted by ν, a convention followed here.

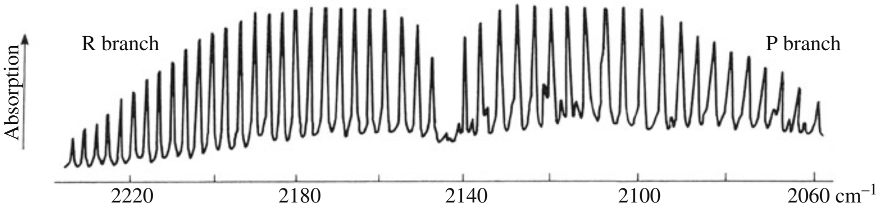


Fig. 3.2 Vibration-rotation band of carbon monoxide

At this resolution the CO molecule appears as an oscillator: as discussed in the energy loss experiment of Chap. 1, Sect. 1.4, the quantum number n is sufficient to describe the infrared spectrum of diatomic molecules. The c.s.c.o. consists of the single operator N or $H_{\text{oscillator}}$, and the basis vectors are labeled by the quantum number $|n\rangle = |E_n\rangle$.

When a spectrometer of sufficiently high resolution is used, as shown in Fig. 3.2, the broad spectral line for the CO molecule around $\nu = 2140\,\text{cm}^{-1}$ is resolved into many almost-equidistant, closely-spaced peaks, with one peak missing in the center of the band. Going out from the gap, there are two branches, the P branch toward longer wavelengths (smaller frequencies) and the R branch toward smaller wavelengths.

This splitting of the broad peak into many narrow peaks cannot be explained by the oscillator model, implying that a state characterized by the quantum number n is not a pure state. Instead it is a mixture of states with different energies. In the oscillator, however, the state characterized by n was a pure state described by a projection operator $\Lambda_n = |n\rangle\langle n|$ on a one-dimensional subspace spanned by $|n\rangle$, namely $\Lambda_n \mathscr{H}$. From the fine structure in the spectral line of Fig. 3.2, it follows that the state of the diatomic molecule characterized by the vibrating quantum number n has split into sublevels. Consequently the oscillator model alone describes only part of the properties of the diatomic molecule. To describe the finer details of the spectrum, the oscillator model must be combined with the rotator model with the result that the c.s.c.o. consists of the set of three operators

$$\{H_{\text{oscillator}}, H_{\text{rotator}}, J_3\} \text{ or } \{N, \mathbf{J}^2, J_3\}\,. \tag{3.3.22}$$

Had the initial resolution been with high precision at lower energies, the spectral lines of the rotator in the far infrared (10^{-3} eV) of Fig. 3.2 would have been detected first. The c.s.c.o. for the rotator consists of the two operators $\mathbf{J}^2$ and J_3. Increasing the energy to 10^{-1} eV, the vibrational degrees of the spectra can be observed because the distance r between the mass points is no longer constant but, instead, is given by the operator $\sqrt{\mathbf{Q}^2}$ with the eigenvectors $|x\rangle$. For the harmonic oscillator, instead of using the eigenvectors $|x\rangle$, it is possible to use the eigenvectors $|n\rangle$ of the operator $H_{\text{oscillator}} = \hbar\omega(N + \frac{1}{2}1)$ and the c.s.c.o. is again given by (3.3.22).

As stated in Fundamental Postulate I, a quantum-physical system is described by an operator-algebra of observables although for particular physical questions, physicists often only work with subalgebras. From the experimental data a physicist determines how many quantum numbers are required and the possible values of these quantum numbers. According to the Fundamental Postulate III, this information allows the determination of the c.s.c.o. $\{A_k\}$ and its spectrum. The total algebra $\mathscr{A}$ is then conjectured by adding to $\{A_k\}$ a minimum number of other operators such that the matrix elements of $\mathscr{A}$ calculated from the properties of this algebra agree with the experimental values of the corresponding observables.

If an experiment yields more values than can be supplied by a given number of commuting operators, then the system is not complete: one or more new quantum numbers must be introduced, and the system of commuting operators must be correspondingly enlarged. Finally the algebra itself must also usually be further enlarged by adding additional operators. For example, the operators $J_x = J_1$ and $J_y = J_2$ must be added to the c.s.c.o. (3.3.22) such that the J_1, J_2, J_3 fulfill the commutation relations of SO(3) and, together with the c.s.c.o., an algebra of observables is formed that correctly describes additional features of a physical system such as the vibrating-rotator.

As illustrated by the vibrating-rotator model, an increase in either energy or precision can necessitate an enlargement of the c.s.c.o. and the algebra of observables. Increasing the energy acting on a rotating molecule introduces the vibrational degrees of freedom and necessitates the introduction of the operators a and $a^\dagger$. Increasing the precision of the energy resolution reveals the rotational degrees of freedom of the dumbbell and requires the introduction of the angular momentum operators J_i.

Because of the procedure described above for obtaining the algebra of observables $\mathscr{A}$, it is assumed that any algebra describing a physical system always has a c.s.c.o. Thus for the algebra $\mathscr{A}$ there exists a c.s.c.o. ,

$$A_1, A_2, \ldots, A_N \in \mathscr{A}\,, \tag{3.3.23}$$

that have a set of (generalized) eigenvectors,[4]

$$A_k|\lambda_1, \ldots, \lambda_N\rangle = \lambda_k|\lambda_1, \ldots, \lambda_N\rangle\,, \tag{3.3.24}$$

such that every physical state vector ϕ can be represented as follows:

$$\phi = \int_\Lambda d\mu\langle\lambda)|\lambda_1, \ldots, \lambda_N\rangle\langle\lambda_1, \ldots, \lambda_N|\phi\rangle\,. \tag{3.3.25}$$

---

[4] Mathematically this assumption can be stated: the conditions of the nuclear spectral theorem are fulfilled.

Here the set $\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ is the spectrum of the c.s.c.o. (3.3.23) with $\{\lambda_k\}$ being the spectrum of the observable A_k. If A_k has a continuous spectrum, then $\{\lambda_k\}$ is a set of generalized eigenvalues; if A_k has a discrete spectrum, then $\{\lambda_k\}$ is the set of all discrete eigenvalues; and if A_k has a discrete and continuous spectrum, then $\{\lambda_k\}$ is the set of all discrete and continuous eigenvalues. For the latter case, the symbol $\int_\Lambda d\mu(\lambda) = \int_{\Lambda_1} \int_{\Lambda_2} \cdots \int_{\Lambda_N} d\mu(\lambda_1, \ldots, \lambda_N)$ in (3.3.25) indicates a summation over the discrete spectrum and integration over the continuous spectrum. Thus (3.3.25) is the generalization of the spectral theorem from the single variable to N variables $\lambda_1, \ldots, \lambda_N$.

It often happens that some of the observables $A_1, A_2, \ldots, A_M$ of the c.s.c.o. have only a continuous spectrum $\Lambda_C = \{(\lambda_1, \ldots, \lambda_M)\}$ and the other observables $A_{M+1}, \ldots, A_N$ of the c.s.c.o. have only a discrete spectrum $\Lambda_D = \{(\lambda_{M+1}, \ldots, \lambda_N)\}$. Then (3.3.25) takes the form

$$\phi = \int_{\Lambda_C} \cdots \int d\mu(\lambda_1, \lambda_2, \ldots, \lambda_M) \sum_{\Lambda_D} \cdots \sum |\lambda_1, \ldots, \lambda_M, \lambda_{M+1}, \ldots, \lambda_N\rangle \times \langle \lambda_1, \ldots, \lambda_N | \phi \rangle \,, \qquad (3.3.26)$$

where the integration extends over a set of generalized eigenvalues $\{\lambda_k\}$ with $k = 1, 2, \ldots, M$ of all observables $A_1, \ldots, A_M$, and the summation runs over all discrete eigenvalues $\{\lambda_k\}$, $k = M+1, M+2, \ldots, N$ of all observables $A_{M+1}, \ldots, A_N$, and the statement (3.3.25) or (3.3.26) is the basis of the Dirac formulation of quantum mechanics.

In the precise mathematical formulation, using the Rigged Hilbert Space,[5] (3.3.26) is the general nuclear spectral theorem for a complete system of commuting operators with absolutely continuous spectrum. (Operators that have a singularly continuous spectrum are excluded, but there has been no need for such operators in quantum physics.)

A right-hand side

$$\Phi \subset \mathscr{H} \subset \Phi^x \qquad (3.3.27)$$

is constructed such that the algebra of observables is represented by an algebra of *continuous* operators $\mathscr{A}$ in a Schwartz space Φ. The conjugate operators $A^x \in \mathscr{A}^x$ in the dual space of Φ are defined using

$$\langle \phi | A^x F \rangle = \langle A\phi | F \rangle \ \text{ for all } \ \phi \in \Phi,\ F \in \Phi^x \,. \qquad (3.3.28)$$

$\mathscr{A}^x$ also forms an algebra of operators, and they are continuous operators in the space of continuous functionals Φ^x. The nuclear spectral theorem then assures that

---

[5] A. Bohm, *The Rigged Hilbert Space and Quantum Mechanics*, Lecture Notes in Physics, **78** (1978), Springer-Verlag, Berlin, Heidelberg, New York.

for a *complete set of commuting observables*

$$A_1, A_2, \dots, A_N \quad [A_j, A_k] = 0 \tag{3.3.29}$$

there is a system of generalized eigenvectors

$$A_i|\lambda_1, \lambda_2, \dots, \lambda_N\rangle = \lambda_i \; |\lambda_1, \lambda_2, \dots, \lambda_N\rangle \tag{3.3.30}$$

such that every $\phi \in \Phi$ can be expanded with respect to this basis system as in (3.3.26).

The spectral theorem for nuclear spaces (3.3.26) thus provides the mathematical foundation for Dirac's formalism provided that the convergence in Φ can be defined such that the nuclear spectral theorem is fulfilled. That is, as long as Φ is a nuclear space.

The standard example of an algebra of operators is the enveloping algebra $\mathscr{E}(G)$ of the Lie group G represented by continuous operators in the space Φ.[6] In order to represent the elements of the Lie algebra (and therewith the elements of the enveloping algebra $\mathscr{E}(G)$) by an algebra of *continuous* operators in a space Φ, the convergence in the space Φ must be defined by the following countable set of scalar products:

$$(\psi, \phi)_p = (\psi, (\Delta + 1)^p \phi, \quad \|\phi\|_p = \sqrt{(\phi, \phi)_p}\,; \quad \|\phi\|_0 \le \|\phi\|_1 \le \|\phi\|_2 \cdots \tag{3.3.31}$$

where $p = 1, 2, 3, \dots$. Here Δ is the Nelson operator defined by $\Delta = \sum X_i^2$ where X_i, $i = 1, 2, 3, \dots, N$ are the generators of the unitary representation of G and are also the generators of the enveloping algebra $\mathscr{E}(G)$. For example, for $\mathscr{E}(SO(3))$, $\Delta = \mathbf{J}^2$, and for $\mathscr{E}(E(3))$, $\Delta = \mathbf{J}^2 + \mathbf{Q}^2$. The convergence of a sequence of $\phi_\nu \in \Phi$ is defined by

$$\phi_\nu \xrightarrow{\tau_\Phi} \phi \quad \Leftrightarrow \quad \|\phi_\nu - \phi\|_p \to 0 \; \text{ for every } p = 0, 1, 2, 3, \dots . \tag{3.3.32}$$

The Nelson operator Δ is essentially self-adjoint in Φ iff the X_i are generators of a unitary representation and the X_i are continuous operators with respect to convergence in the space Φ. Possible applications of this mathematical result in physics are the Spectrum Generating Groups G or the Symmetry Groups S described by unitary group representations $g \to U(g)$ in $\mathscr{H}$.

---

[6] For example, the group G could be the rotation group $SO(3)$ describing an elementary rotator and $\mathscr{E}(SO(3))$ is the algebra generated by the angular momentum operators J_i discussed in Chap. 2, Sect. 2.3. Or if G is the group $E(3)$, then $\mathscr{E}(G)$ is the algebra generated by the momentum operators J_i and the position operators Q_i describing the rotating dumbbell of Chap. 2, Sect. 2.4.

For a self-adjoint Hamiltonian, which is often an element of $\mathscr{E}(G) = \mathscr{E}(3)$, this leads to generalized eigenvectors of the Hamiltonian H:

$$|E, j, j_3\rangle \in \Phi^{x} \quad \text{with} \quad \langle H\phi|E, j, j_3\rangle = E\langle\phi|E, j, j_3\rangle \text{ for all } \phi \in \Phi, \quad 0 \leq E < \infty\,, \tag{3.3.33}$$

and to time evolution for the observable $|\psi\rangle\langle\psi|$ is given by the unitary group $U(t) = e^{iHt/\hbar}$:

$$\langle e^{iHt/\hbar}\psi|E, j, j_3\rangle = \langle\psi|e^{iHt/\hbar}|E, j, j_3\rangle = e^{-iEt/\hbar}\,\langle\psi|E, j, j_3\rangle\,, \quad -\infty < t < \infty\,, \tag{3.3.34}$$

where $E = E_j = \hbar^2\, j(j+1)/(2I)$ in the case of the rigid rotator $H_{\text{rotator}} = \mathbf{J}^2/(2I)$.

An immediate consequence of the statement (3.3.24), (3.3.25), or (3.3.26) is the following: Let $A_1, \ldots, A_N$ be a c.s.c.o. with (generalized) eigenvectors $|\lambda_1, \ldots, \lambda_N\rangle = |\lambda\rangle$ such that

$$A_k|\lambda\rangle = \lambda_k|\lambda\rangle \quad (k = 1, 2, \ldots, N)\,. \tag{3.3.35}$$

Then if another system of eigenvectors $|a>= |a_1, a_2, \ldots, a_N>$ is found with the property

$$A_i|a>= a_i|a> \qquad (i = 1, 2, \ldots, N) \tag{3.3.36}$$

it follows that

$$\{a_i\} = \{\lambda_i\} \text{ and } |a>= \alpha|\lambda\rangle\,, \tag{3.3.37}$$

where $\alpha \in \mathbb{C}$.

3.4 The Vibrating Rotator

The vibrating rotator is an instructive example for the combination of two quantum physical systems that has clearly been observed in the infrared spectra of diatomic molecules. A theoretical description requires the combination of the oscillator (Chap. 1) and the rotator (Chap. 2).

The space of physical states of the oscillator is denoted $\mathscr{H}_{\text{oscillator}}$, and a basis system of eigenvectors of the operator $H_{\text{oscillator}}$ in $\mathscr{H}_{\text{oscillator}}$ is

$$\text{basis system for } \mathscr{H}_{\text{oscillator}} : \; |n\rangle, \quad n = 0, 1, 2, \ldots\,. \tag{3.4.1}$$

The action of all observables (all elements of the algebra of observables) of the quantum-mechanical oscillator on the basis vectors $|n\rangle$ is known from Chap. 1.

Similarly, in Chap. 2 the space of physical states of the rotator is denoted $\Re$. In $\Re$ a basis system of eigenvectors of the operators L_3 and $\mathbf{L}^2$ is

$$\text{basis system for } \Re : |\ell, m\rangle, \quad l = 0, 1, 2, \ldots, \quad m = -\ell, -\ell + 1, \ldots, \ell\,. \tag{3.4.2}$$

The above basis vectors are eigenvectors of the operators L_3 and $H_{\text{rotator}} = \mathbf{L}^2/(2I)$ with respective eigenvalues $\hbar m$ and $\hbar^2 \ell(\ell + 1)/(2I)$.

The space of physical states of the vibrating rotator is, according to Fundamental Postulate IV, the direct-product space

$$\mathscr{H}_{\text{vibrating-rotator}} = \mathscr{H}_{\text{oscillator}} \otimes \Re\,, \tag{3.4.3}$$

and the observables are the operators $\sum_i A^{(i)}_{\text{oscillator}} \otimes A^{(j)}_{\text{rotator}}$ where $A^{(i)}_{\text{oscillator}}$ is any observable of the oscillator and $A^{(j)}_{\text{rotator}}$ is any observable of the rotator. The basis system in $\mathscr{H}_{\text{vibrating-rotator}}$ is the direct product of the basis systems $|n\rangle$ in $\mathscr{H}_{\text{oscillator}}$ and $|\ell, m\rangle$ in $\Re$, and is denoted by $|n, \ell, m\rangle$,

$$|n, \ell, m\rangle = |n\rangle \otimes |\ell, m\rangle\,. \tag{3.4.4}$$

The classical picture for such a vibrating rotator is a dumbbell consisting of two mass points (atomic nuclei) connected by a spring with spring constant k as shown in Fig. 3.3.

For the idealized situation in which the harmonic vibrations do not affect the rigid rotation and vice versa, the energy operator of this physical combination of the harmonic oscillator and rigid rotator is given by

$$H = H_{\text{oscillator}} \otimes 1 + 1 \otimes H_{\text{rotator}}\,, \tag{3.4.5}$$

where

$$H_{\text{oscillator}} = \frac{\mathbf{P}^2}{2\mu} + \frac{1}{2}kQ^2 = \hbar\omega(N + \frac{1}{2}1)\,, \quad H_{\text{rotator}} = \frac{\mathbf{L}^2}{2I}\,. \tag{3.4.6}$$

In the above equation the constants μ, k and I are, respectively, the reduced mass, the spring constant and the moment of inertia. Treating the molecule as a dumbbell

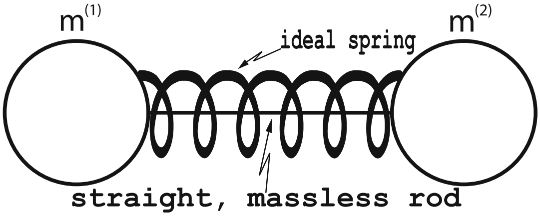


Fig. 3.3 Vibrating, rotating dumbbell

that oscillates and rotates independently is clearly an idealization since oscillations affect the moment of inertia and the rate of rotation affects the equilibrium position of the oscillator. To a first approximation the interplay between oscillations and rotations is neglected with the result that observables have the form,

$$A = A_{\text{oscillator}} \otimes 1 + 1 \otimes A_{\text{rotator}} \,. \tag{3.4.7}$$

The energy spectrum of the idealized vibrating rotator is obtained by applying the Hamiltonian (3.4.5) to the basis vectors (3.4.4),

$$\begin{aligned} H|n, l, m\rangle &= (H_{\text{oscillator}} \otimes 1 + 1 \otimes H_{\text{rotator}})\,(|n\rangle \otimes |\ell, m\rangle) \\ &= (H_{\text{oscillator}}|n\rangle \otimes |\ell, m\rangle + |n\rangle \otimes H_{\text{rotator}}|\ell, m\rangle \\ &= [\hbar\omega\left(n + \frac{1}{2}\right) + \frac{\hbar^2}{2I}\ell\langle\ell + 1)]|n, l, m\rangle \,. \end{aligned} \tag{3.4.8a}$$

The final line in the above equation was obtained by first using (3.4.6) and then (2.3.35). From (3.4.8a) the energy eigenvalues $E_{n,\ell}$ of the vibrating rotator are

$$E_{n,\ell} = \hbar\omega\left(n + \frac{1}{2}\right) + \frac{\hbar^2}{2I}\ell(\ell + 1)\,, \quad n = 0, 1, 2, \ldots\,, \quad \ell = 1, 1, 2, \ldots\,. \tag{3.4.8b}$$

From experiments with vibrating diatomic molecules, the constant $\hbar\omega$ is known to be on the order of 0.2 eV, and from rotational spectra, $\hbar^2/2I$ is on the order of 10^{-3} eV. Thus pure vibrational and pure rotational transitions are, respectively, in the near and far infrared. The energy levels of a vibrating rotator are shown in Fig. 3.4 on the next page.

To obtain the frequencies ν of light emitted or absorbed as a transition occurs between energy levels, from Chap. 1 recall that harmonic oscillator transitions occur between adjacent levels. Similarly, dipole transitions of the rigid rotator occur between adjacent rotator levels. Thus the selection rules for the vibrating rotator are as follows:

$$\text{harmonic oscillator:} \qquad \Delta n = \pm 1 \tag{3.4.9a}$$

$$\text{rotator:} \qquad \Delta \ell = \pm 1 \tag{3.4.9b}$$

If $\Delta n = 0$, then only the rotational spectrum resulting from rotation is observed, and if $\Delta \ell = 0$, then only the spectrum resulting from oscillation is observed. The more interesting spectra occur when both n and ℓ change. For a harmonic oscillator transition from n to $n + 1$ (energy absorption), the two possibilities for which there is a rotator transition are $\Delta \ell = \pm 1$. From (3.4.8b), the frequencies of absorbed light

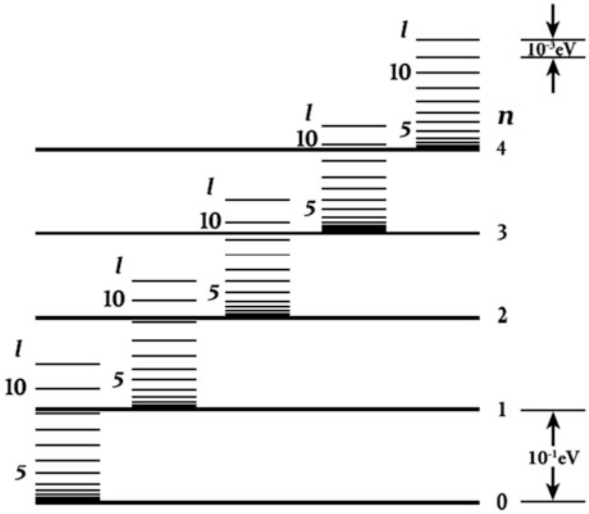


Fig. 3.4 Energy levels of the vibrating rotator. For each of the first five vibrational levels (long horizontal lines), a number of rotational levels (short horizontal lines) are drawn [from Herzberg (1966), with permission]

are as follows:

$$\nu_R = \frac{E_{n+1,\ell+1} - E_{n,\ell}}{2\pi\hbar} = \frac{1}{2\pi}\left[\omega + \frac{\hbar(\ell+1)}{I}\right] \qquad \Delta\ell = +1, \tag{3.4.10a}$$

$$\nu_P = \frac{E_{n+1,\ell-1} - E_{n,\ell}}{2\pi\hbar} = \frac{1}{2\pi}\left[\omega - \frac{\hbar\ell}{I}\right] \qquad \Delta\ell = -1. \tag{3.4.10b}$$

Transitions for which $\Delta\ell = 1$ and $\Delta\ell = -1$ are called, respectively, the R branch and the P branch.

The energy levels of a diatomic molecule can be determined by a spectrometer that detects photons absorbed (or emitted) as molecules make transitions from one energy level to another. If the resolution (or sensitivity) of the spectrometer is not sufficiently good, a single broad spectral line for CO at $\nu = 2140\ \text{cm}^{-1}$ is observed that corresponds to transitions between adjacent harmonic oscillator levels.[7] However, if the resolution of the spectrometer is improved, the broad spectral line for CO is resolved into a whole band of narrow spectral lines that occur as a result of the rotation of the molecule and are shown in Fig. 3.2 on page 148. Figure 3.5 shows the transitions between the energy bands of two adjacent vibrational states and their correspondence to the spectra of Fig. 3.2. The branch

---

[7] As is customary in the literature on molecular spectra, frequencies have been converted to wave-number units by dividing by a factor of c (in cm/s), resulting in frequencies measured in cm^{-1}.

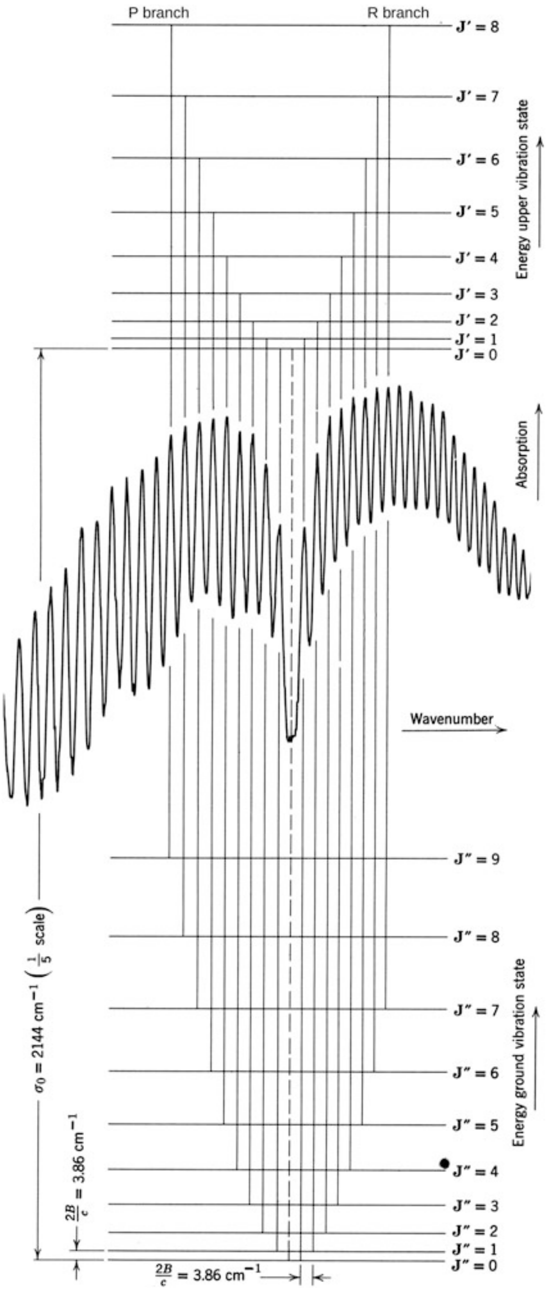


Fig. 3.5 Origin and appearance of rotational structure. P and R branches are shown to the left and right, respectively, on the spectrometer tracing of the CO fundamental absorption band at 2144 cm^{-1}. The Q branch (dashed line) is missing. Energy levels are shown to scale except that the distance between upper and lower vibrational states (2144 cm^{-1} should be about five times as great as in the figure [from Bauman (1962), with permission]

with increasing frequencies is called the R branch. For the R branch the minimum value of ℓ is zero corresponding to a transition from $\ell = 0$ to $\ell = 1$. From (3.4.10a), the R branch frequencies are

$$\nu_R = (\omega + \hbar/I)/2\pi,\ (\omega + 2\hbar/I)/2\pi,\ (\omega + 3\hbar/I)/2\pi, \dots . \tag{3.4.11a}$$

The branch with decreasing frequencies is called the P branch, and the minimum value of ℓ is unity, corresponding to a transition from $\ell = 1$ to $\ell = 0$. From (3.4.10b), these frequencies are

$$\nu_P = (\omega - \hbar/I)/2\pi,\ (\omega - 2\hbar/I)/2\pi,\ (\omega - 3\hbar/I)/2\pi, \dots . \tag{3.4.11b}$$

Note that as predicted by the spectra for ν_R and ν_P in (3.4.11), the frequency $\omega/(2\pi)$ is missing! In Fig. 3.2 on page 148 a corresponding "spike" is missing at $\omega/(2\pi) = \nu = 2140\ \text{cm}^{-1}$.

From the experimental values for the gaps between successive spikes of the vibrational-rotational spectra of the CO molecule depicted in Fig. 3.2, the following experimental values are determined:

$$\frac{\hbar}{2\pi I_{\text{CO}}} = 1.18 \times 10^{11}\ \text{Hz}, \tag{3.4.12a}$$

or

$$\frac{\hbar}{2\pi c I_{\text{CO}}} = 3.94\ \text{cm}^{-1}. \tag{3.4.12b}$$

The relationship between an absorbed frequency and the pair of energy levels between which the transition took place are shown in Fig. 3.5 on the facing page. Comparing with (3.4.10), the formulas for ν_R and ν_P (in cm^{-1}) are, respectively,

$$\nu_R = \frac{E_{n+1,\ell+1} - E_{n,\ell}}{2\pi\hbar c} = \nu_0 + 2B(\ell + 1), \qquad \Delta\ell = +1, \tag{3.4.13a}$$

$$\nu_P = \frac{E_{n+1,\ell-1} - E_{n,\ell}}{2\pi\hbar c} = \nu_0 - 2B\ell, \qquad \Delta\ell = -1, \tag{3.4.13b}$$

where

$$\nu_0 = \frac{\omega}{2\pi c}, \qquad B = \frac{\hbar}{4\pi c I}. \tag{3.4.14}$$

Frequencies have been converted to wave-number units in the above two formulas by dividing by a factor of c. The frequency spectrum calculated in (3.4.11) or (3.4.13) is depicted in strip (b) of Fig. 3.6 on the next page.

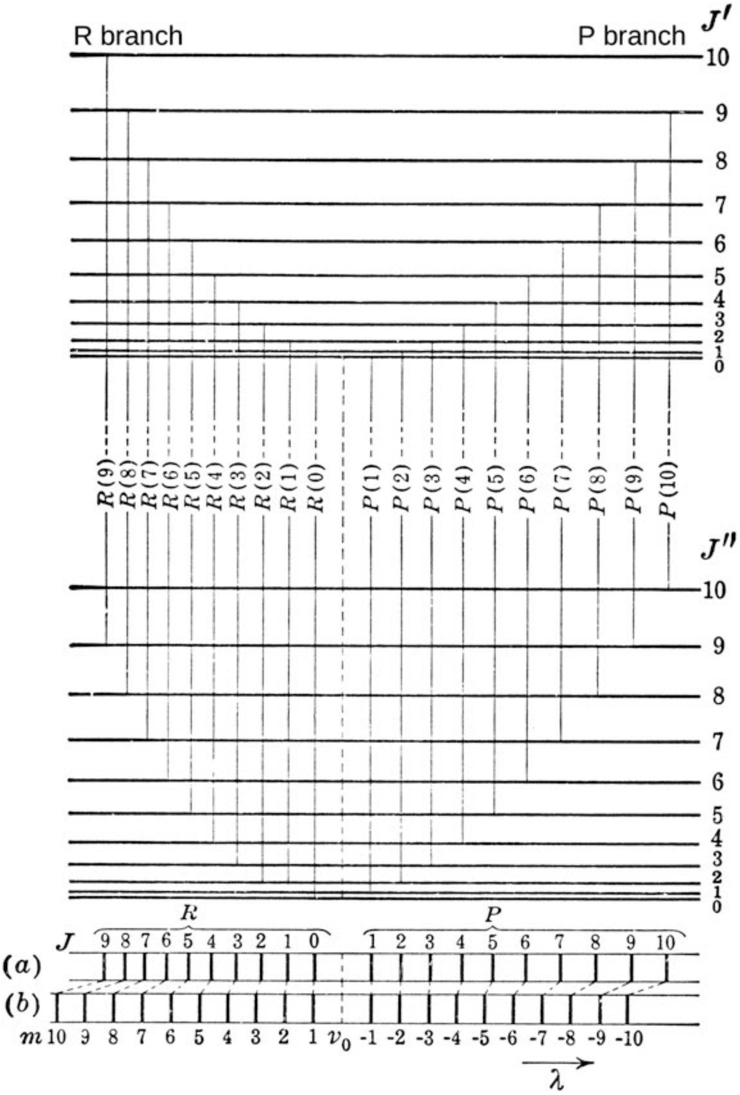


Fig. 3.6 Energy-level diagram explaining the fine structure of a rotation-vibration band. In general the separation of the two vibrational levels is considerably larger compared with the spacing of the rotational levels than shown in the figure (indicated by the broken parts of the vertical lines representing the transitions). The schematic spectrograms (**a**) and (**b**) give the resulting spectrum with and without allowance for the interaction between rotation and vibration. In these spectrograms, unlike most of the others, short wavelengths are at the left. Note that the R- and P-branches are on opposite sides of Figs. 3.5 on page 156 and 3.6 [from Herzberg (1966), with permission]

3.4.1 The Interplay of Rotations and Vibrations

The observed spectrum from Fig. 3.5 on page 156 is depicted in strip (a) of Fig. 3.6 on the preceding page. The prediction of the vibrating-rotator model without interaction between vibration and rotation agrees well, but not precisely. Because a vibrating rotator cannot be rigid, the moment of inertia I cannot be a constant. As a result of this interplay between rotations and vibrations, the observed lines in the R branch draw closer together, and those in the P branch move farther apart than indicated by the equidistant lines.

Since the moment of inertia I is different in different vibrational states, I depends on the vibrational quantum number, $I \to I_n$. Instead of (3.4.8b)

$$E_{nl} = \hbar\omega(n + \frac{1}{2}) + \frac{1}{2I_n}\hbar^2\ell(\ell + 1) \tag{3.4.15}$$

The moment of inertia is no longer a constant system parameter but instead is an operator that takes different expectation values depending on the vibrational state, $I_n = \langle n|I|n\rangle$.

With the aid of (3.4.15), the wave numbers of the resulting lines are

$$\nu = \frac{E_{n'\ell'} - E_{n''\ell''}}{2\pi\hbar c} = \nu_0(n' - n'') + B_{n'}\ell'(\ell' + 1) - B_{n''}\ell''(\ell'' + 1)\,, \tag{3.4.16a}$$

where

$$B_n = \frac{h}{8\pi^2 c I_n}\,. \tag{3.4.16b}$$

The absorption frequencies (in wave number units cm^{-1}) for transitions $n''l'' \to n'l'$ between neighboring vibrational levels are as follows:

$$\nu_R = \nu_0(n' - n'') + 2B_{n'} + (3B_{n'} - B_{n''})\ell + (B_{n'} - B_{n''})\ell^2$$
$$(\ell' = \ell + 1,\ \ell'' = \ell,\ \Delta\ell = +1)\,, \tag{3.4.17a}$$

$$\nu_P = \nu_0(n' - n'') - (B_{n'} + B_{n''})\ell + (B_{n'} - B_{n''})\ell^2$$
$$(\ell' = \ell - 1,\ \ell'' = \ell,\ \Delta l = -1)\,. \tag{3.4.17b}$$

Equations (3.4.17) give an excellent agreement with the empirical fine structure of the infrared bands.

For the HCl molecule, the values of B_n have been determined empirically for the various bands $n' \leftrightarrow n''$:

$$0 \leftrightarrow 1 \quad 0 \leftrightarrow 2 \quad 0 \leftrightarrow 3 \quad 0 \leftrightarrow 4 \quad 0 \leftrightarrow 5 \tag{3.4.18}$$

Table 3.3 Rotational constants of HCl in the different vibrational levels of the electronic ground state [from Herzberg (1966), p. 800, with permission]

n	B_n (cm^{-1})	ΔB_n (cm^{-1})
0	10.4400	0.3034
1	10.1366	0.3037
2	9.8329	0.2986
3	9.5343	0.302
4	9.232	0.299
5	8.933	

Transitions with $\Delta n > 1$ occur as a consequence of the small anharmonicity, and the values of B_n are summarized in Table 3.3. As the value of the vibrational quantum number increases, the amplitude of vibration increases, the moment of inertia increases and, from (3.4.16b), B_n decreases. The difference $\Delta B_n \equiv B_n - B_{n+1}$ between successive values is very nearly a constant, so that B_n can be fitted by the formula

$$B_n = B_e - \alpha_e(n + \frac{1}{2}), \tag{3.4.19}$$

where α_e is a constant small compared with $B_e = 10.5909$ cm^{-1}, the equilibrium value of B_n.

Although experimental values for the vibration-rotation spectra of the CO molecule are neither as numerous nor as accurate as those for HCl, from the spectrum depicted in Fig. 3.5 on page 156 it is possible to calculate

$$B^{\mathrm{CO}} = 1.96 \text{ cm}^{-1}. \tag{3.4.20}$$

This value, and hence the fine structure in the energy spectrum, is considerably smaller than that for HCl as given in Table 3.3.

The diatomic molecule with the largest rotational constant B_e, and thus the largest energy difference between rotational levels, is the H_2 molecule for which $B_e^{H_2} = 60.80\,\text{cm}^{-1}$.

A qualitative theoretical explanation of (3.4.15), where the empirical expression for I_n can be obtained by combining (3.4.16b) and (3.4.19), follows from the classical picture Fig. 3.3 on page 153 of the diatomic molecule as two rigid spheres connected by a spring. When the dumbbell is in a state of higher vibrational energy, it has a larger amplitude and consequently a larger moment of inertia, implying that I_n^{-1} decreases with increasing n.

For the quantum-mechanical observables, the empirical formula (3.4.19) reveals that the form (3.4.5) is inadequate to describe the energy levels of a diatomic molecule: an interaction Hamiltonian H_{int} between vibrational and rotational degrees of freedom is needed. The preceding discussion suggests an interaction term of the form

$$H_{\text{int}} = g H_{\text{oscillator}} \otimes H_{\text{rotator}}, \tag{3.4.21}$$

where g is a coupling constant of dimension $(\mathrm{eV})^{-1}$. Including the interaction Hamiltonian (3.4.21), the energy operator for the vibrating, rotating, interacting diatomic molecule is given by

$$H = H_{\text{oscillator}} \otimes 1 + 1 \otimes H_{\text{rotator}} + gH_{\text{oscillator}} \otimes H_{\text{rotator}}\,, \tag{3.4.22a}$$

where

$$H_{\text{oscillator}} = \hbar\omega(N + \frac{1}{2}1)\,, \tag{3.4.22b}$$

and

$$H_{\text{rotator}} = \frac{1}{2I_e}\mathbf{L}^2\,. \tag{3.4.22c}$$

The characteristic angular frequency $\omega = \sqrt{k/\mu}$ is the system constant of the oscillator, and I_e is the value of the moment of inertia at the equilibrium separation x_e. That is, $I_e = \mu x_e^2$ is the system constant of the rotator.

The energy values of the diatomic molecule with vibration-rotation interaction are the expectation values of H in the physical states. It is not obvious that the vectors $|n, \ell, m\rangle = |n\rangle \otimes |\ell, m\rangle$ of (3.4.4), where $|n\rangle$ are eigenvectors of N and $|\ell, m\rangle$ are eigenvectors of $\mathbf{L}^2$ and L_3, represent the pure states of this physical system. However, as they also happen to be eigenstates of the energy operator H of (3.4.22a), they are the obvious choice for states in an energy measurement. The energy values are the eigenvalues of H in the basis $|n, \ell, m\rangle$:

$$E_{nl} = \hbar\omega(n + \frac{1}{2}) + \frac{1}{2I_e}\hbar^2\ell(\ell+1) + g\hbar\omega\frac{\hbar^2}{2I_e}(n + \frac{1}{2})\ell(\ell+1)\,. \tag{3.4.23}$$

The wave number $\nu_{n\ell}$ (frequency in cm^{-1}) of the radiation quantum corresponding to the energy value $E_{n\ell}$ is

$$\nu_{n\ell} = \frac{E_{n\ell}}{2\pi\hbar c}\,, \tag{3.4.24}$$

and is called the *term value*.[8]

From (3.4.23) the term values of the vibrating rotator are

$$\nu_{n\ell} = \nu_0(n + \frac{1}{2}) + [B_e - \alpha_e(n + \frac{1}{2})]\ell(\ell+1)\,, \tag{3.4.25}$$

---

[8] Note that the same symbol ν is used for the frequency (in s^{-1}) and the wave number (in cm^{-1}).

where the following standard notation of molecular spectroscopy has been used:

$$B_e = \frac{\hbar}{4\pi c I_e} = \frac{\hbar}{4\pi c \mu x_e^2}\,, \qquad \alpha_e = -\frac{g\nu_0\hbar^2}{2I_e}\,, \qquad B_n = B_e - \alpha_e(n+\frac{1}{2})\,. \tag{3.4.26}$$

According to the above qualitative considerations, I_n^{-1} (and, therefore, B_n) depends on n. As a consequence, α_e should be larger than zero, a condition that is always fulfilled experimentally. Equation (3.4.25) with (3.4.26) gives the following wave numbers for transitions in the R branch:

$$\nu_R = \nu_{n'\ell+1} - \nu_{n''\ell} = \nu_0(n'-n'') + 2B_{n'} + (3B_{n'} - B_{n''})\ell + (B_{n'} - B_{n''})\ell^2\,. \tag{3.4.27a}$$

The wave numbers in the P branch are

$$\nu_P = \nu_{n'\ell-1} - \nu_{n''\ell} = \nu_0(n'-n'') - (B_{n'} + B_{n''})\ell + (B_{n'} - B_{n''})\ell^2\,. \tag{3.4.27b}$$

The above two formulas are, respectively, the well-established empirical formulas (3.4.17a) and (3.4.17b), verifying that (3.4.21) was a good guess.

Equation (3.4.25) is not the end of the story of the vibrating and rotating diatomic molecule because anharmonic effects of the oscillator and the influence of centrifugal forces have not been taken into account. When these effects are also taken into account, to a higher degree of accuracy the term values of a vibrating rotator are

$$\nu_{n\ell} = \omega_e(n+\frac{1}{2}) - \omega_e\xi_e(n+\frac{1}{2})^2 + B_n\ell(\ell+1) - D_n\ell^2(\ell+1)^2\,. \tag{3.4.28}$$

In the above equation ω_e is the standard notation for

$$\omega_e = \frac{\omega}{2\pi c} = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}} = \nu_0\,, \tag{3.4.29a}$$

where ν_0 was originally defined in Eq. (3.4.14). ξ_e is a small parameter expressing the anharmonicity,[9] and

$$B_n = B_e - \alpha_e(n+\frac{1}{2})\,, \tag{3.4.29b}$$

$$D_n = D_e + \beta_e(n+\frac{1}{2})\,. \tag{3.4.29c}$$

B_e is given by (3.4.26).

---

[9] $\omega_e\xi_e$ is given by the anharmonicity of the oscillator (terms proportional to Q^3, Q^4, etc.) and can be calculated by perturbation theory.

According to the semiclassical consideration leading to (V.4.56), D_e may be expressed in terms of the reduced mass μ, equilibrium separation x_e and spring constant k:

$$D_e = \frac{\hbar^3}{4\pi c k \mu^2 x_e^6} \,. \tag{3.4.30}$$

From (3.4.29a), (3.4.26) and (3.4.30) it follows that the three system parameters D_e, B_e and ω_e are not independent but instead are related by

$$D_e = \frac{4B_e^3}{\omega_e^2} \,. \tag{3.4.31}$$

The parameters ξ_e, α_e and β_e that express the degree of anharmonicity are known empirically to be small:

$$\xi_e \ll 1 \,, \quad \frac{\alpha_e}{B_e} \ll 1 \,, \quad \frac{\beta_e}{D_e} \ll 1 \,, \tag{3.4.32}$$

which must be the case since they represent the effect of corrections to models that are rather precisely realized by microphysical systems in nature. The system parameters ω_e, ξ_e, B_e, D_e, α_e and β_e have been experimentally determined for many diatomic molecules and are collected in tables [cf. Herzberg (1966)]. Eq. (3.4.28) gives a very good description of the vibration-rotation spectra of diatomic molecules, and higher-order corrections are needed only in exceptional cases.

Diatomic molecules are vibrating rotators only as long as the internal energy is sufficiently low-roughly, in the region of energy of infrared radiation. For higher energies in the range 1–20 eV, corresponding to the visible and ultraviolet regions, the molecules are no longer just vibrating rotators because new degrees of freedom become accessible to electronic transitions. However, in each electronic state the molecule is still a vibrating rotator. This leads to energy spectra depicted schematically in Fig. 3.7 on the next page for two electronic states. The electronic structure of molecules will not be discussed here, but it is similar to the electronic structure of atoms.

3.5 Addition of Two Angular Momenta: Clebsch-Gordan Coefficients

In classical mechanics angular momentum is a vector quantity, implying that two angular momenta $\mathbf{S}_a^{\text{classical}}$ and $\mathbf{S}_b^{\text{classical}}$ are added vectorially, $\mathbf{S}_{\text{total}}^{\text{classical}} = \mathbf{S}_a^{\text{classical}} + \mathbf{S}_b^{\text{classical}}$.

From the correspondence principle, if quantum system a has an angular momentum operator $\mathbf{S}_a$ in the space $\Re_a$, and if a second quantum system b has an angular momentum operator $\mathbf{S}_b$ in the space $\Re_b$, then combining the two systems according to Fundamental Postulate IV, the total angular momentum operator $\mathbf{S}$ of

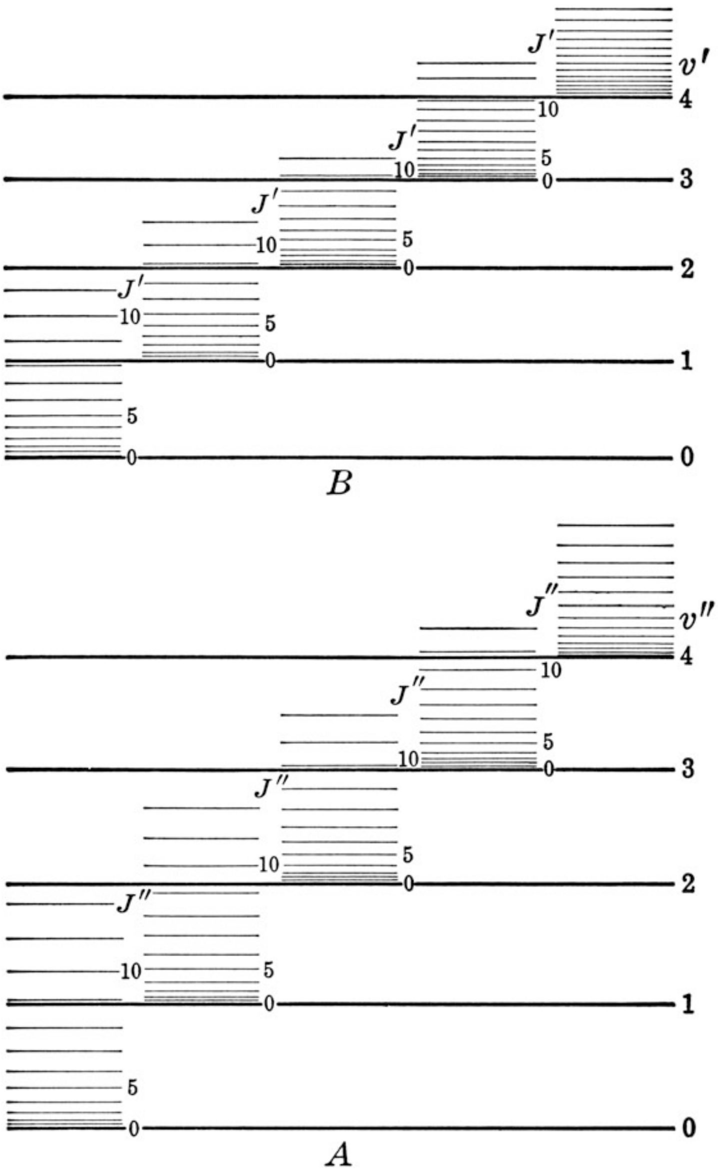


Fig. 3.7 Vibrational and rotational levels of two electronic states A and B of a molecule (schematic). Only the first few rotational and vibrational levels are drawn in each case [from Herzberg (1966), with permission]

the combined system is

$$\mathbf{S} = \underbrace{\mathbf{S}_a \times 1}_{\mathbf{S}_a} + \underbrace{1 \times \mathbf{S}_b}_{\mathbf{S}_b} \tag{3.5.1}$$

in the direct-product space

$$\Re = \Re_a \otimes \Re_b = \Re^{j_a} \otimes \Re^{j_b}\,. \tag{3.5.2}$$

There are many physical examples for such a combination of two quantum physical rotators. For example, a deuteron is a bound state of a proton and a neutron, each of which has a spin = 1/2. If the proton and neutron have no relative orbital angular momentum, the total angular momentum $\mathbf{S}_D$ of the deuteron is the sum of the spin angular momenta of the two constituents, $\mathbf{S}_D = \mathbf{S}_p + \mathbf{S}_n$. For the hydrogen atom, the total angular momentum **J** of the orbiting electron is the sum of its orbital angular momentum **L** and its spin angular momentum **S**,

$$\mathbf{J} = \mathbf{L} + \mathbf{S}\,. \tag{3.5.3}$$

When only the rotational properties of an elementary particle such as a proton, neutron, or electron are considered, the system is called an *elementary rotator*. Here the proton in the deuteron is taken to be an elementary rotator with spin angular momentum $s_p = 1/2$; its space of physical states is therefore the space $\Re_p^{s=1/2}$ of angular momentum s=1/2. Similarly, the neutron is also an elementary rotator with $s_n = 1/2$ in a space of physical states $\Re_n^{s=1/2}$. Because all other properties of the neutron and proton are being ignored, these two spaces are the same (isomorphic) angular momentum spaces $\Re^{1/2}$ of Chap. 2. To an excellent approximation, the neutron and proton in a deuteron have no orbital angular momentum. Thus, according to the Fundamental Postulate IV, the rotational properties of the deuteron are described by the direct-product space[10]

$$\Re = \Re_p^{s=1/2} \otimes \Re_n^{s=1/2}\,. \tag{3.5.4}$$

Similarly, for the orbiting and spinning electron in the hydrogen atom with orbital angular momentum ℓ and spin-1/2 the space of physical states is

$$\Re = \Re_m^{\ell} \otimes \Re^{1/2}\,. \tag{3.5.5}$$

Equations (3.5.4) and (3.5.5) are special cases of a very general result. In accordance with Fundamental Postulate IV for combining physical systems, whenever two rotational motions are combined, one described by $\Re_a^j$ and the other by $\Re_b^j$, the combined rotational motion is described by the direct-product space

$$\Re = \Re^{j_a} \otimes \Re^{j_b}\,. \tag{3.5.6}$$

---

[10] For an orbital angular momentum $\ell = 2$ component of the deuteron see H. Frauenfelder and E. M. Henley, Subatomic Physics, Sect. 14.5, Prentice Hall (1991).

Equation (3.5.6) is true unless $\mathfrak{R}^{j_a}$ and $\mathfrak{R}^{j_b}$ describe identical quantum systems. If $j_a = j_b$ are half-integers, the constituents are fermions, and the anti-symmetric subspace of (3.5.6) must be chosen. If $j_a = j_b$ are integers, the constituents are bosons, and the symmetric subspace of (3.5.6) must be chosen. For the total angular momentum of an electron in a hydrogen atom, the orbital angular momentum ℓ ($j_a = \ell$) of the electron is combined with the spin $s_e = 1/2$ ($j_b = 1/2$) of the electron, so j_a and j_b are associated with the same constituent. In the deuteron j_a and j_b are, respectively, the spin of the proton and neutron so they are associated with different constituents. There are numerous other examples in quantum physics for the combination—also called "addition"—of two angular momenta.

The general case described by (3.5.6) will now be discussed. The two angular momenta that are to be combined are denoted $\mathbf{J}_a$ and $\mathbf{J}_b$. Each may represent either spin, orbital angular momentum, or both. The operator $\mathbf{J}_a$ acts in the space $\mathfrak{R}_a^j$ and $\mathbf{J}_b$ acts in the space $\mathfrak{R}_b^j$.

Classically, if an object possesses both orbital angular momentum $\mathbf{l}$ and spin angular momentum $\mathbf{s}$, the total angular momentum $\mathbf{j}$ is the vector sum $\mathbf{j} = \mathbf{l} + \mathbf{s}$. From the correspondence between classical and quantum mechanics, the total angular momentum operator $\mathbf{J}$ is given by

$$\mathbf{J} = \mathbf{J}_a \otimes 1 + 1 \otimes \mathbf{J}_b \,, \tag{3.5.7}$$

where $\mathbf{J}_a$ acts in $\mathfrak{R}^{j_a}$ and $\mathbf{J}_b$ acts in $\mathfrak{R}^{j_b}$. In component notation the above equation is

$$J_i = J_{ai} \otimes 1 + 1 \otimes J_{bi} \,, \quad i = 1, 2, 3 \;\text{ or }\; i = x, y, z \,. \tag{3.5.8}$$

Experiments have verified that angular momentum in quantum mechanics is additive as expressed by (3.5.6), which are special cases of Fundamental Postulate IV.

Before delving into the details of how to add angular momentum in quantum mechanics, first consider the problem from a classical viewpoint. If a deuteron is made from a proton and neutron with respective spins $s_p = 1/2$ and $s_n = 1/2$ and no relative orbital angular momentum, it is reasonable to guess that if the spins are "parallel," the total angular momentum $j = 1/2 + 1/2 = 1$, and if the spins are "antiparallel", the total angular momentum is $1/2 - 1/2 = 0$. This intuitive result is correct. A surprise occurs, however, when a vector diagram is drawn corresponding to the operator relationship $\mathbf{J} = \mathbf{S}_p + \mathbf{S}_n$. Since the eigenvalue of $\mathbf{J}^2$ is $\hbar^2 j(j+1)$, the "length" of $\mathbf{J}$ is $\hbar\sqrt{j(j+1)}$, which equals $\hbar\sqrt{2}$ for $j = 1$. Similarly, for $s = 1/2$, the "length" of $\mathbf{S}_p$ and $\mathbf{S}_n$ is $\hbar\sqrt{3}/2$. From Fig. 3.8 on the facing page it is immediately apparent that the expectation value of the spin operators $\mathbf{S}_p$ and $\mathbf{S}_n$ are not parallel. This quantum effect occurs because the expectation value of $\mathbf{J}^2$ is $\hbar^2 j(j+1)$, not $\hbar^2 j^2$. From the above discussion, it follows that adding angular momenta in quantum mechanics is more involved than just adding vectors.

From the algebraic commutation relations (2.3.1), the properties of the operators J_{ai} and J_{bi} were determined. In particular, for each integer or half-integer value of

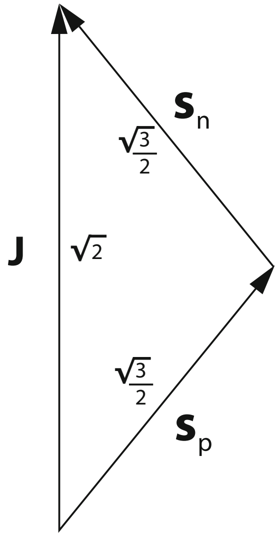


Fig. 3.8 Vector relationship $\mathbf{J} = \mathbf{S}_p + \mathbf{S}_n$ for j=1, s_p=1/2 and s_n=1/2

j_a or j_b, there is a representation space $\mathfrak{R}^{j_a}$ or $\mathfrak{R}^{j_b}$, respectively, in which the J_{ai} and J_{bi} act. Since the operators J_{ai} and J_{bi} act in different spaces, they commute, $[J_{ai}\,,\,J_{bj}] = 0$. It then immediately follows that the total angular momentum operator $\mathbf{J}$ as given in (3.5.7) fulfills the commutation relations (2.3.1). Therefore, the J_i act in the direct-product space (3.5.6) that is characterized by a single number j, called the total angular momentum, which is either an integer or half integer. The following question then immediately arises: in the direct-product space (3.5.6), which values of the angular momentum are possible for $\mathbf{J}$ of (3.5.7)? In other words, when two physical systems with respective angular momenta j_a and j_b are combined, what are the possible values j of total angular momentum? In the remainder of this section this and related questions will be answered.

The basis vectors of $\mathfrak{R}^{j_a}$ and $\mathfrak{R}^{j_b}$, denoted respectively by $|j_a, m_a\rangle$ and $|j_b, m_b\rangle$, satisfy

$$\mathbf{J}_a^2|j_a, m_a\rangle = \hbar^2 j_a(j_a+1)|j_a, m_a\rangle\,, \quad J_{a3}|j_a, m_a\rangle = \hbar m_a|j_a, m_a\rangle\,, \tag{3.5.9a}$$

$$\mathbf{J}_b^2|j_b, m_b\rangle = \hbar^2 j_b(j_b+1)|j_b, m_b\rangle\,, \quad J_{b3}|j_b, m_b\rangle = \hbar m_b|j_b, m_b\rangle\,. \tag{3.5.9b}$$

A basis system in the direct-product space $\mathfrak{R} = \mathfrak{R}^{j_a} \otimes \mathfrak{R}^{j_b}$ is, according to the definition of $\mathfrak{R}$, given by

$$|j_a, m_a\rangle \otimes |j_b, m_b\rangle \equiv |j_a, m_a, j_b, m_b\rangle\,, \tag{3.5.10}$$

and is called the direct-product basis. The left-hand side of (3.5.10) is the definition of basis vectors in the direct-product space (3.5.6), and the right-hand side defines a new symbol for the basis vectors.

In the direct-product space $\Re$, using vector notation the angular momentum operators are given by (3.5.7), and using component notation are given by (3.5.8). The basis system (3.5.10) consists of eigenstates of the following complete set of commuting operators

$$\mathbf{J}_a^2,\, J_{a3},\, \mathbf{J}_b^2,\, J_{b3}\,, \tag{3.5.11}$$

with the respective eigenvalues

$$\mathbf{J}_a^2|j_a, m_a\rangle \otimes |j_b, m_b\rangle = \hbar^2 j_a(j_a+1)|j_a, m_a\rangle \otimes |j_b, m_b\rangle\,, \tag{3.5.12a}$$

$$J_{a3}|j_a, m_a\rangle \otimes |j_b, m_b\rangle = \hbar m_a|j_a, m_a\rangle \otimes |j_b, m_b\rangle\,, \tag{3.5.12b}$$

$$\mathbf{J}_b^2|j_a, m_a\rangle \otimes |j_b, m_b\rangle = \hbar^2 j_b(j_b+1)|j_a, m_a\rangle \otimes |j_b, m_b\rangle\,, \tag{3.5.12c}$$

$$J_{b3}|j_a, m_a\rangle \otimes |j_b, m_b\rangle = \hbar m_b|j_a, m_a\rangle \otimes |j_b, m_b\rangle\,. \tag{3.5.12d}$$

A basis is said to be physical if it consists of eigenstates in which the physical system appears, and the basis (3.5.10) is, in general, not a physical basis for a combined system. Stationary states (states that do not change in time) are always eigenstates—or mixtures of eigenstates—of the Hamiltonian H, so physical systems such as atoms and molecules are eigenstates of the Hamiltonian. The state vectors that describe these physical states must consist of eigenvectors of operators that commute with the Hamiltonian. If the physical system is spherically symmetric or rotationally invariant, implying that $[H, J_i] = 0$, then not all of the J_{ai},and J_{bi} commute with the Hamiltonian, and the direct-product eigenvectors (3.5.10) are in general not eigenvectors of the Hamiltonian and, as a consequence, are not a physical basis system.

A rotationally invariant Hamiltonian commutes with the total angular momentum $\mathbf{J}^2$. Energy eigenstates can, therefore, be eigenstates of total angular momentum $\mathbf{J}^2$ and J_3 where

$$\begin{aligned}\mathbf{J}^2 = (\mathbf{J}_a + \mathbf{J}_b)\cdot(\mathbf{J}_a + \mathbf{J}_b) &= \mathbf{J}_a^2 + \mathbf{J}_b^2 + 2\sum_{i=1}^{3} J_{ai}J_{bi}\\ &= \mathbf{J}_a^2 + \mathbf{J}_b^2 + 2J_{a3}J_{b3} + J_{a+}J_{b-} + J_{a-}J_{b+}\,,\end{aligned} \tag{3.5.13a}$$

$$J_3 = J_{a3} + J_{b3}\,. \tag{3.5.13b}$$

But energy eigenstates are usually not eigenvectors of J_{a3} and J_{b3}.

The four eigenvalues of the four operators (3.5.11) are required to completely specify the basis system (3.5.10). Therefore any other complete set of commuting operators will usually consist of four operators provided only discrete eigenvalues are involved. While $\mathbf{J}^2$ and J_3 don't represent a complete set of commuting operators, it is easy to determine two additional operators that commute with $\mathbf{J}^2$ and J_3. Since $[J_{ai}, J_{bk}] = 0$ and $\mathbf{J}_a^2$ commutes with all J_{ai}, $\mathbf{J}_a^2$ also commutes with $\mathbf{J}^2$ and J_3. Similarly, $\mathbf{J}_b^2$ commutes with $\mathbf{J}^2$ and J_3. Since the operators $\mathbf{J}_a^2$ and $\mathbf{J}_b^2$

commute with each other, a complete set of commuting operators is also given by

$$\mathbf{J}_a^2,\ \mathbf{J}_b^2,\ \mathbf{J}^2,\ J_3\,. \tag{3.5.14}$$

A basis system of normalized eigenvectors for these operators is denoted by

$$|j_a, j_b, j, m\rangle\,. \tag{3.5.15}$$

These total angular momentum eigenvectors satisfy

$$\mathbf{J}_a^2|j_a, j_b, j, m\rangle = \hbar^2 j_a(j_a+1)|j_a, j_b, j, m\rangle\,, \tag{3.5.16a}$$

$$\mathbf{J}_b^2|j_a, j_b, j, m\rangle = \hbar^2 j_b(j_b+1)|j_a, j_b, j, m\rangle\,, \tag{3.5.16b}$$

$$\mathbf{J}^2|j_a, j_b, j, m\rangle = \hbar^2 j(j+1)|j_a, j_b, j, m\rangle\,, \tag{3.5.16c}$$

$$J_3|j_a, j_b, j, m\rangle = \hbar m|j_a, j_b, j, m\rangle \tag{3.5.16d}$$

and are called the total angular momentum basis vectors. Often the respective eigenvalues $\hbar^2 j_a(j_a+1)$ and $\hbar^2 j_b(j_b+1)$ of $\mathbf{J}_a^2$ and $\mathbf{J}_b^2$ are fixed so the labels j_a and j_b are omitted for notational simplicity, and $|j_a, j_b, j, m\rangle \equiv |j, m\rangle$.

The two different basis systems (3.5.10) and (3.5.15) span the same space $\mathfrak{R}$ in (3.5.6); consequently, each of the basis vectors (3.5.10) can be expressed in terms of the the basis system (3.5.15),

$$|j_a, m_a, j_b, m_b\rangle = \sum_{j,m} |j_a, j_b, j, m\rangle\langle j_a, j_b, j, m|j_a, m_a, j_b, m_b\rangle\,, \tag{3.5.17}$$

and each basis vector of the basis (3.5.15) can be expanded with respect to the basis system (3.5.10),

$$|j_a, j_b, j, m\rangle = \sum_{m_a, m_b} |j_a, m_a, j_b, m_b\rangle\langle j_a, m_a, j_b, m_b|j_a, j_b, j, m\rangle\,. \tag{3.5.18}$$

The expansion coefficients $\langle j_a, m_a, j_b, m_b|j_a, j_b, j, m\rangle$ are the scalar product of the vector $|j_a, m_a\rangle \otimes |j_b, m_b\rangle$ with $|j_a, j_b, j, m\rangle$ and are called *Clebsch-Gordan* or *Wigner coefficients*. For fixed values of j_a and j_b they are denoted in various ways in the literature, the most common being

$$\begin{aligned}\langle j_a, m_a, j_b, m_b|j_a, j_b, j, m\rangle &= (\langle j_a, m_a| \otimes \langle j_b, m_b|, |j_a, j_b, j, m\rangle)\,,\\ &= \langle j_a, m_a, j_b, m_b|j, m\rangle\,,\\ &= C(j_a, j_b, j, m_a, m_b, m)\,,\\ &= C^{j_a\ j_b\ j}_{m_a\, m_b\, m}\,,\\ &= \langle m_a, m_b|j, m\rangle\,.\end{aligned} \tag{3.5.19}$$

In the final expression for Clebsch-Gordan coefficients in the above equation, the quantum numbers j_a and j_b have been suppressed because they remain unchanged.

Taking the scalar product of (3.5.17) with $|j_a, m'_a, j_b, m'_b\rangle$ the orthogonality relation of the Clebsch-Gordan coefficients is obtained:

$$\begin{aligned}\delta_{m'_a,m_a}\delta_{m'_b,m_b} &= \langle j_a, m'_a, j_b, m'_b|j_a, m_a, j_b, m_b\rangle \\ &= \sum_{j,m}\langle j_a, m'_a, j_b, m'_b|j_a, j_b, j, m\rangle\langle j_a, j_b, j, m|j_a, m_a, j_b, m_b\rangle\,. \end{aligned} \tag{3.5.20}$$

Similarly, taking the scalar product of (3.5.18) with $|j_a, j_b, j', m'\rangle$, a second orthogonality relation relation results:

$$\begin{aligned}\delta_{j,j'}\delta_{m,m'} &= \langle j_a, j_b, j', m'|j_a, j_b, j, m\rangle \\ &= \sum_{m_a,m_b}\langle j_a, j_b, j', m'|j_a, m_a, j_b, m_b\rangle\langle j_a, m_a, j_b, m_b|j_a, j_b, j, m\rangle\,. \end{aligned} \tag{3.5.21}$$

The Clebsch-Gordan coefficients are needed when two quantum systems with j_a and j_b are combined into a single quantum system with "total" angular momentum j. This is an important problem in physics: if the system is spherically symmetric or rotationally invariant, then the eigenvectors $|j_a, j_b, j, m\rangle$ of total angular momentum can be energy eigenstates, whereas the direct-product basis vectors are usually not. Thus the transformation (3.5.18) is very important because it relates the direct-product basis to the physical basis, and a knowledge of the transformation coefficients (3.5.19) permits this transformation to be made.

Several properties of the space $\mathfrak{R} = \mathfrak{R}^{j_a} \otimes \mathfrak{R}^{j_b}$ of (3.5.6) can readily be determined: (1) How many eigenvectors are required to span the space $\mathfrak{R}$? For the basis system (3.5.10), there are $(2j_a + 1)$ eigenvectors $|j_a, m_a\rangle$ and $(2j_b + 1)$ eigenvectors $|j_b, m_b\rangle$. Therefore, there are $(2j_a + 1) \times (2j_b + 1)$ linearly independent eigenvectors $|j_a, m_a\rangle \otimes |j_b, m_b\rangle$. But since the eigenvectors (3.5.15) span the same space, there must be $(2j_a + 1) \times (2j_b + 1)$ linearly independent eigenvectors $|j_a, j_b, j, m\rangle$. (2) What are the possible values of j in the space $\mathfrak{R}$? The basis vectors in the basis system (3.5.8) are also eigenvectors of J_3,

$$\begin{aligned}J_3|j_a, m_a\rangle \otimes |j_b, m_b\rangle &= (J_{a3} + J_{b3})|j_a, m_a\rangle \otimes |j_b, m_b\rangle \\ &= [J_{a3}|j_a, m_a\rangle] \otimes |j_b, m_b\rangle + |j_a, m_a\rangle \otimes [J_{b3}|j_b, m_b\rangle] \\ &= \hbar(m_a + m_b)|j_a, m_a\rangle \otimes |j_b, m_b\rangle\,, \end{aligned} \tag{3.5.22}$$

with an eigenvalue $m = m_a + m_b$. Since the maximum value of m_a is j_a, and the maximum value of m_b is j_b, the maximum value of m is $j_a + j_b$. Therefore, the maximum value of the total angular momentum j is $j_{\max} = j_a + j_b$. This result is easy to understand intuitively: the maximum value of total angular momentum occurs when $\mathbf{J}_a$ and $\mathbf{J}_b$ are "parallel". Similarly, if $\mathbf{J}_a$ and $\mathbf{J}_b$ are "antiparallel", the total angular momentum would be a minimum and is given by $j_{\min} = |j_a - j_b|$. The maximum and minimum values for j are now known. What other values are allowed? Since the eigenvalue m of J_3 changes in integer steps from $m_{\max} = j_a + j_b$, the only allowed values of j must differ from $j_{\max}$ by an integer. From the discussion it is not obvious that every value of j that can appear, will actually appear and that each value of j will appear only once. However, it turns out that each value of j does appear exactly once,[11] so the possible values of j are

$$j = j_a + j_b,\ j_a + j_b - 1,\ j_a + j_b - 2,\ \ldots,\ |j_a - j_b|\,. \tag{3.5.23}$$

The above result can be written in terms of spaces as

$$\mathfrak{R} = \mathfrak{R}^{j_a} \otimes \mathfrak{R}^{j_b} = \mathfrak{R}^{j_a+j_b} \oplus \mathfrak{R}^{j_a+j_b-1} \oplus \cdots \oplus \mathfrak{R}^{|j_a-j_b|}\,, \tag{3.5.24}$$

and is called the reduction of the direct-product space into a direct sum of irreducible total angular momentum spaces.

Summarizing, the space $\mathfrak{R}^{j_a} \otimes \mathfrak{R}^{j_b}$ is, in general, not an irreducible representation space or ladder representation space of the algebra of total angular momentum $\mathscr{E}(\mathrm{SO}(3)_{j_i})$. That is, not all vectors of $\mathfrak{R}^{j_a} \otimes \mathfrak{R}^{j_b}$ can be obtained by applying the $J_\pm$ a sufficient number of times to one of its vectors. Instead, the space $\mathfrak{R}^{j_a} \otimes \mathfrak{R}^{j_b}$ is is the direct sum of several such irreducible representation spaces $\mathfrak{R}^{j}$ as given in (3.5.24). Also, the Clebsch-Gordan coefficients, which are the transition coefficients between the two basis systems (3.5.10) and (3.5.15), are zero unless $m = m_a + m_b$ and j is given by one of the values in (3.5.23):

$$C^{j_a\ j_b\ j}_{m_a\,m_b\,m} = 0 \ \text{ for } \ m \neq m_a + m_b,\quad j \neq j_a + j_b, j_a + j_b - 1, \ldots, |j_a - j_b|\,. \tag{3.5.25}$$

Before deriving a general formula for Clebsch-Gordan coefficients, to familiarize the reader with the ideas and mathematics, Clebsch-Gordan coefficients are first derived for two specific, simple examples.

Example 3.5.1 Calculate the Clebsch-Gordan coefficients for the deuteron, treated as a combination of a proton with spin−1/2 ($j_a = 1/2$) and a neutron with spin−1/2 ($j_b = 1/2$) and no relative orbital angular momentum.

Solution In the direct-product basis (3.5.8), where $\mathbf{J}_a^2$, J_{a3}, $\mathbf{J}_b^2$, and J_{b3} are a complete set of commuting operators, there are four independent eigenvectors,

$$\begin{aligned} &\left|\frac{1}{2},\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},\frac{1}{2}\right\rangle, \quad \left|\frac{1}{2},\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},-\frac{1}{2}\right\rangle, \\ &\left|\frac{1}{2},-\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},\frac{1}{2}\right\rangle, \quad \left|\frac{1}{2},-\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},-\frac{1}{2}\right\rangle, \end{aligned} \tag{3.5.26}$$

---

[11] M.E. Rose, *Elementary Theory of Angular Momentum* (New York: John Wiley, 1957).

corresponding to $m_a = \pm 1/2$ and $m_b = \pm 1/2$, respectively. Since $j_a + j_b = 1$ and $j_a - j_b = 0$, according to (3.5.23) three states of total angular momentum $j = 1$ are expected with $m = 1, 0$, and -1. In addition, one state of total angular momentum $j = 0$ is expected.

Letting $\mathbf{J}^2$ act on the basis system (3.5.10) and using (3.5.13a),

$$\begin{aligned}\mathbf{J}^2|j_a, m_a\rangle \otimes |j_b, m_b\rangle &= [\mathbf{J}_a^2|j_a, m_a\rangle] \otimes |j_b, m_b\rangle + |j_a, m_a\rangle \otimes [\mathbf{J}_b^2|j_b, m_b\rangle] \\ &+ 2[J_{a3}|j_a, m_a\rangle] \otimes [J_{b3}|j_b, m_b\rangle] + [J_{a+}|j_a, m_a\rangle] \otimes [J_{b-}|jb, m_b\rangle] \\ &+ [J_{a-}|j_a, m_a\rangle] \otimes [J_{b+}|j_b, m_b\rangle]\,. \end{aligned} \tag{3.5.27}$$

From (2.3.43b) and (2.3.43c), the respective action of the raising and lowering operator on the vector $|\ell, m\rangle$ is

$$L_{\pm}|\ell, m\rangle = \hbar\sqrt{(\ell \mp m)(\ell \pm m + 1)}|\ell, m \pm 1\rangle\,, \tag{3.5.28a}$$

where

$$L_{\pm} = J_{\pm},\ J_{a\pm}, \text{or}\ \ J_{b\pm}\,. \tag{3.5.28b}$$

Using (3.5.28a), (3.5.27) becomes,

$$\begin{aligned}\mathbf{J}^2|j_a, m_a\rangle \otimes |j_b, m_b\rangle &= \hbar^2[j_a(j_a + 1) + j_b\langle j_b + 1) + 2m_a m_b]|j_a, m_a\rangle \otimes |j_b, m_b\rangle \\ &+ \hbar^2\sqrt{(j_a - m_a)(j_a + m_a + 1)(j_b + m_b)\langle j_b - m_b + 1)}|j_a, m_a + 1\rangle \otimes |j_b, m_b - 1\rangle \\ &+ \hbar^2\sqrt{(j_a + m_a)(j_a - m_a + 1)(j_b - m_b)\langle j_b + m_b + 1)}|j_a, m_a - 1\rangle \otimes |j_b, m_b + 1\rangle. \end{aligned} \tag{3.5.29}$$

From (3.5.29) it immediately follows that

$$\mathbf{J}^2\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \hbar^2 1(1 + 1)\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle, \tag{3.5.30a}$$

$$\mathbf{J}^2\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \hbar^2 1(1 + 1)\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle, \tag{3.5.30b}$$

which verifies that these two states are eigenstates of total angular momentum j = 1. Actually this was already known before carrying out the calculation in (3.5.30). From the preceding discussion, only spin-1 and spin-0 states are present. Using (3.5.22), the eigenvectors $|1/2, 1/2\rangle \otimes |1/2, 1/2\rangle$ and $|1/2, -1/2\rangle \otimes |1/2, -1/2\rangle$ have $m = 1$ and $m = -1$, respectively, so these two states must be spin-1 states. In terms of the basis system (3.5.15) (up to a phase factor),

$$\left|j_a = \frac{1}{2}, j_b = \frac{1}{2}, j = 1, m = 1\right\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle, \tag{3.5.31a}$$

$$\left|j_a = \frac{1}{2}, j_b = \frac{1}{2}, j = 1, m = -1\right\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle, \tag{3.5.31b}$$

The only remaining problem is to find the linear combinations of the two states with $m = 0$, namely $|1/2, 1/2\rangle \otimes |1/2, -1/2\rangle$ and $|1/2, -1/2 \otimes |1/2, 1/2\rangle$ that have $j = 1$ and $j = 0$, respectively. This calculation will be done in two different ways. The first method is straightforward and not the least bit elegant. The second is almost as straightforward and more elegant.

Consider a linear combination of the two states with $m = 0$, which is an eigenvector with angular momentum j, where j equals either 0 or 1.

$$\left|\frac{1}{2}, \frac{1}{2} j, 0\right\rangle = \alpha\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle + \beta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle. \tag{3.5.32}$$

Here α and β are constants. Allowing $\mathbf{J}^2$ to operate on $|1/2, 1/2, j, 0\rangle$ and using (3.5.29),

$$\mathbf{J}^2\left|\frac{1}{2}, \frac{1}{2} j, 0\right\rangle = \hbar^2(\alpha + \beta)\left[\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle + \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\right], \tag{3.5.33a}$$

But since $|1/2, 1/2, j, 0\rangle$ is an eigenstate of $\mathbf{J}^2$ with angular momentum j,

$$\begin{aligned} \mathbf{J}^2\left|\frac{1}{2}, \frac{1}{2} j, 0\right\rangle &\equiv \hbar^2 j(j+1)\left|\frac{1}{2}, \frac{1}{2}, j, 0\right\rangle \\ &= \hbar^2 j(j+1)\left[\alpha\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle + \beta\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\right]. \end{aligned} \tag{3.5.33b}$$

Equating (3.5.33a) and (3.5.33b), and choosing $j = 0$, requires $\alpha = -\beta$. To normalize $|1/2, 1/2, 0, 0\rangle$, the choice $\alpha = -\beta = 1/\sqrt{2}$ is made with the result,

$$\left|\frac{1}{2}, \frac{1}{2}, 0, 0\right\rangle = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle - \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\right]. \tag{3.5.34}$$

Repeating the above calculation with $j = 1$ yields the condition $\alpha = \beta$. Again choosing $\alpha = 1/\sqrt{2}$ for purposes of normalization,

$$\left|\frac{1}{2}, \frac{1}{2}, 1, 0\right\rangle = \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle + \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\right]. \tag{3.5.35}$$

The states (3.5.34) and (3.5.35) are orthogonal as required. Note that the eigenstates with $j = 1$ are symmetric under the interchange $a \leftrightarrows b$ while the $j = 0$ state is antisymmetric.

A more elegant method for determining the state $|1/2, 1/2, 1, 0\rangle$ is to begin with the eigenstate $|1/2, 1/2, 1, 1\rangle$ in (3.5.31a) and then let the total angular momentum lowering operator $J_- = (J_{a1} + J_{b1}) - i(J_{a2} + J_{b2}) = J_{a-} + J_{b-}$ act on it to create the desired state $|1/2, 1/2, 1, 0\rangle$. From (3.5.28a),

$$J_-\left|\frac{1}{2}, \frac{1}{2}, 1, 1\right\rangle = \hbar\sqrt{2}\left|\frac{1}{2}, \frac{1}{2}, 1, 0\right\rangle. \tag{3.5.36}$$

Therefore, from (3.5.36),

$$\begin{aligned}\left|\frac{1}{2}, \frac{1}{2} 1, 0\right\rangle &= \frac{1}{\hbar\sqrt{2}} J_-\left|\frac{1}{2}, \frac{1}{2}, 1, 1\right\rangle \\ &= \frac{1}{\hbar\sqrt{2}}\left\{\left[J_{a-}\left|\frac{1}{2}, \frac{1}{2}\right\rangle\right] \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left[J_{b-}\left|\frac{1}{2}, \frac{1}{2}\right\rangle\right]\right\} \\ &= \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle\right],\end{aligned} \tag{3.5.37}$$

which agrees with (3.5.35). The spin-0 state is just the $m = 0$ state that is orthogonal to (3.5.35) and can be found accordingly. Alternatively the eigenstate $|1/2, 1/2, 1, 0\rangle$ could have been found by applying the raising operator J_+ to the eigenstate $|1/2, 1/2, 1, -1\rangle$ in (3.5.31b).

Example 3.5.2 Calculate the matrix element $M = \langle 1/2, 1/2, 1, 0|\mathbf{J}_a \cdot \mathbf{J}_b|1/2, 1/2, 1, 0\rangle$.

Solution With the aid of (3.5.13a), the matrix element can be written as

$$M = \left\langle\frac{1}{2}, \frac{1}{2}, 1, 0\right|\frac{1}{2}\left(\mathbf{J}^2 - \mathbf{J}_a^2 - \mathbf{J}_b^2\right)\left|\frac{1}{2}, \frac{1}{2}, 1, 0\right\rangle.$$

Now $|1/2, 1/2, 1, 0\rangle$ is an eigenstate of each of the operators $\mathbf{J}^2$, $\mathbf{J}_a^2$, and $\mathbf{J}_b^2$ as shown in (3.5.16). Thus

$$M = \left\langle\frac{1}{2}, \frac{1}{2}, 1, 0\right|\frac{\hbar^2}{2}\left[1(1+1) - \frac{1}{2}\left(\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right)\right]\left|\frac{1}{2}, \frac{1}{2}, 1, 0\right\rangle = \frac{\hbar^2}{4}.$$

For the case of the deuteron, which was just considered, the two-dimensional spaces spanned by $|j_a = 1/2, m_a = 1/2\rangle$ and $|j_b = 1/2, m_b = 1/2\rangle$ are denoted, respectively, by $\Re^{j_a=1/2}$ and $\Re^{j_b=1/2}$. The direct-product space spanned by $|1/2, 1/2, j, m\rangle$ is denoted $\Re$. To indicate that $\Re$ is the direct-product space, $\Re$ is written as

$$\Re = \Re^{j_a=1/2} \otimes \Re^{j_b=1/2} . \tag{3.5.38}$$

Because the space $\Re$ contains a spin-1 subspace, denoted $\Re^1$, and a spin-0 subspace, denoted $\Re^0$, the space $\Re$ is expressed as

$$\Re = \Re^{ja=1/2} \otimes \Re^{jb=1/2} = \Re^1 \oplus \Re^0 . \tag{3.5.39}$$

The above equation is called a reduction of $\Re$ into a sum of irreducible angular momentum spaces and is a special case of the general result (3.5.24).

To determine the eigenstates of total angular momentum, the basis system (3.5.15) is expressed in terms of the basis system (3.5.10). From (3.5.18) and (3.5.19) this expansion takes the form

$$|j_a, j_b, j, m\rangle = \sum_{m_a, m_b} C^{j_a\ j_b\ j}_{m_a\ m_b\ m} |j_a, m_a\rangle \otimes |j_b, m_b\rangle . \tag{3.5.40}$$

For the case $j_a = j_b = 1/2$, the Clebsch-Gordan or Wigner coefficients have just been calculated. Comparing (3.5.40) with (3.5.31), (3.5.34) and (3.5.35),

$$\begin{aligned} &C^{1/2\ 1/2\ 1}_{1/2\ 1/2\ 1} = 1 , && C^{1/2\ \ 1/2\ \ 1}_{-1/2\ -1/2\ -1} = 1 , && C^{1/2\ \ 1/2\ \ 0}_{1/2\ -1/2\ 0} = \frac{1}{\sqrt{2}} , \\ &C^{1/2\ 1/2\ 0}_{-1/2\ 1/2\ 0} = \frac{-1}{\sqrt{2}} , && C^{1/2\ \ 1/2\ \ 1}_{1/2\ -1/2\ 0} = \frac{1}{\sqrt{2}} , && C^{1/2\ 1/2\ 1}_{-1/2\ 1/2\ 0} = \frac{1}{\sqrt{2}} . \end{aligned} \tag{3.5.41}$$

Example 3.5.3 Calculate the Clebsch-Gordan coefficients for an electron (spin $\equiv j_b = 1/2$) orbiting a spinless nucleus with orbital angular momentum $\ell \equiv j_a$ (an integer).

Solution From Eq. (3.5.23) the two possible values of angular momentum are $j_a + j_b = \ell + 1/2$ and $j_a - j_b = \ell - 1/2$. From (3.5.22) all eigenstates of J_3 which have an eigenvalue m must be of the form

$$\left|\ell, \frac{1}{2}, j, m\right\rangle = \alpha \left|\ell, m - \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \beta \left|\ell, m + \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle . \tag{3.5.42}$$

The constants α and β must now be determined such that (3.5.42) is an eigenstate of total angular momentum with either $j = \ell + 1/2$ or $j = \ell - 1/2$. This, of course, is the same procedure used to determine α and β in (3.5.32). Letting $\mathbf{J}^2$ operate on (3.5.42) and using (3.5.29) to simplify the right-hand side,

$$\begin{aligned}
\mathbf{J}^2\left|\ell, \frac{1}{2}, j, m\right\rangle &= \hbar^2\left\{\alpha\left[\ell(\ell+1) + \frac{3}{4} + \left(m - \frac{1}{2}\right)\right]\right.\\
&\qquad \left. + \beta\sqrt{(\ell + m + \frac{1}{2})(\ell - m + \frac{1}{2})}\right\}\left|\ell, m - \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\\
&\qquad + \hbar^2\left\{\beta\left[\ell(\ell+1) + \frac{3}{4} - \left(m + \frac{1}{2}\right)\right]\right.\\
&\qquad \left. + \alpha\sqrt{(\ell - m + \frac{1}{2})(\ell + m + \frac{1}{2})}\right\}\left|\ell, m + \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle\\
&= \hbar^2 j(j+1)\left[\alpha\left|\ell, m - \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \beta\left|\ell, m + \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle\right]. \qquad (3.5.43)
\end{aligned}$$

Setting $j = \ell + 1/2$ in (3.5.43) and requiring the coefficients of $|\ell, m - 1/2) \otimes |1/2, 1/2\rangle$ in (3.5.43) be equal, yields the result

$$\alpha = \beta\sqrt{\frac{\ell + m + 1/2}{\ell - m + 1/2}}. \qquad (3.5.44)$$

The same relation is obtained by requiring the coefficients of $|\ell, m + 1/2\rangle \otimes |1/2, -1/2\rangle$ be equal. Normalizing the state such that $\alpha^2 + \beta^2 = 1$ leads to the final result,

$$\begin{aligned}
\left|\ell, \frac{1}{2}, \ell + \frac{1}{2}, m\right\rangle &= \sqrt{\frac{\ell + m + 1/2}{2\ell + 1}}\left|\ell, m - \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\\
&\qquad + \sqrt{\frac{\ell - m + 1/2}{2\ell + 1}}\left|\ell, m + \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle. \qquad (3.5.45)
\end{aligned}$$

The state $|\ell, 1/2, \ell - 1/2, m\rangle$ could be obtained by repeating the above procedure after setting $j = \ell - 1/2$. But $|\ell, 1/2, \ell - 1/2, m\rangle$ is the state with 3-component of angular momentum equal to m that is orthogonal to $|\ell, 1/2, \ell + 1/2, m\rangle$. Therefore,

$$\begin{aligned}
\left|\ell, \frac{1}{2}, \ell - \frac{1}{2}, m\right\rangle &= -\sqrt{\frac{\ell - m + 1/2}{2\ell + 1}}\left|\ell, m - \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\\
&\qquad + \sqrt{\frac{\ell + m + 1/2}{2\ell + 1}}\left|l, m + \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle. \qquad (3.5.46)
\end{aligned}$$

Convenient choices for arbitrary phases have been made in both (3.5.45) and (3.5.46). For example, the minus sign preceding the first term in (3.5.46) could have appeared in front of the second term instead. In the above derivation ℓ has never been required to be an integer. Thus, if ℓ is replaced by j in expressions (3.5.45) and (3.5.46), both equations are true for either integer or half integer values of j.

The Clebsch-Gordan coefficients calculated in this example are given in Table 3.4.

Employing a procedure similar to that used in constructing Table 3.4, Clebsch-Gordan coefficients can be calculated for arbitrary j_a and $j_b = 1$. The results are tabulated in Table 3.5.

Having calculated the Clebsch-Gordan coefficients for two specific, simple cases, $j_a = 1/2$, $j_b = 1/2$ and $j_a = \ell$, $j_b = 1/2$, Clebsch-Gordan coefficients will now be calculated recursively for arbitrary values of j_a, m_a and j_b, m_b using two expressions for the matrix element $\langle j_a, m_a, j_b, m_b | J_\pm | j_a, j_b, j, m\rangle$. Letting $J_\pm$ act to the right and using (3.5.28a),

$$\begin{aligned}\langle j_a, m_a, j_b, m_b | J_\pm | j_a, j_b, j, m\rangle = \\ \hbar\sqrt{(j \mp m)(j \pm m + 1)}\ \langle j_a, m_a, j_b, m_b | j_a, j_b, j, m \pm 1\rangle\,.\end{aligned} \tag{3.5.47a}$$

Since $J_\pm^\dagger = J_\mp = J_{a\mp} + J_{b\mp}$, when $J_\pm$ acts to the left,

$$\langle j_a, m_a, j_b, m_b | J_\pm | j_a, j_b, j, m\rangle = \left((J_{a\mp} + J_{b\mp})|j_a, m_a, j_b, m_b\rangle\,,\ |j_a, j_b, j, m\rangle\right).$$

The second expression for the matrix element is then obtained by again using Eq. (3.5.28a) on page 172:

$$\begin{aligned}&= \hbar\sqrt{(j_a \pm m_a)(j_a \mp m_a + 1)}\,\langle j_a, m_a \mp 1, j_b, m_b | j_a, j_b, j, m\rangle \\ &\quad + \hbar\sqrt{(j_b \pm m_b)(j_b \mp m_b + 1)}\,\langle j_a, m_a, j_b, m_b \mp 1 | j_a, j_b, j, m\rangle\,.\end{aligned} \tag{3.5.47b}$$

Table 3.4 The Clebsch-Gordan coefficients $C^{j_a=\ell}_{m-m_b}{}^{j_b=\frac{1}{2}}_{m_b}{}^{j}_{m}$

	$m_b = \frac{1}{2}$	$m_b = -\frac{1}{2}$
$j = \ell + \frac{1}{2}$	$\sqrt{\frac{\ell+m+1/2}{2\ell+1}}$	$\sqrt{\frac{\ell-m+1/2}{2\ell+1}}$
$j = \ell - \frac{1}{2}$	$-\sqrt{\frac{\ell-m+1/2}{2\ell+1}}$	$\sqrt{\frac{\ell+m+1/2}{2\ell+1}}$

Table 3.5 The Clebsch-Gordan coefficients $C^{j_a}_{m-m_b}{}^{j_b=1}_{m_b}{}^{j}_{m}$

	$m_b = 1$	$m_b = 0$	$m_b = -1$
$j = j_a + 1$	$\sqrt{\frac{(j_a+m)(j_a+m+1)}{(2j_a+1)(2j_a+2)}}$	$\sqrt{\frac{(j_a-m+1)(j_a+m+1)}{(2j_a+1)(j_a+1)}}$	$\sqrt{\frac{(j_a-m)(j_a-m+1)}{(2j_a+1)(2j_a+2)}}$
$j = j_a$	$-\sqrt{\frac{(j_a+m)(j_a-m+1)}{2j_a(j_a+1)}}$	$\frac{m}{\sqrt{j_a(j_a+1)}}$	$\sqrt{\frac{(j_a-m)(j_a+m+1)}{2j_a(j_a+1)}}$
$j = j_a - 1$	$\sqrt{\frac{(j_a-m)(j_a-m+1)}{2j_a(2j_a+1)}}$	$-\sqrt{\frac{(j_a-m)(j_a+m)}{j_a(2j_a+1)}}$	$\sqrt{\frac{(j_a+m+1)(j_a+m)}{2j_a(2j_a+1)}}$

The Clebsch-Gordan coefficients $\langle j_a, m_a, j_b, m_b | j_a, j_b, j, m = j\rangle$ are calculated first by taking $m = j$ in (3.5.47) and calculating the matrix element of J_+. The right-hand side of (3.5.47a) vanishes, and (3.5.47b) becomes

$$\begin{aligned}&\langle j_a, m_a - 1, j_b, m_b | j_a, j_b, j, j\rangle \\ &\qquad = -\left[\frac{(j_b + m_b)(j_b - m_b + 1)}{(j_a + m_a)(j_a - m_a + 1)}\right]^{\frac{1}{2}} \langle j_a, m_a, j_b, m_b - 1 | j_a, j_b, j, j, \rangle \,.\end{aligned} \tag{3.5.48}$$

From (3.5.22) it follows that

$$m = m_a + m_b - 1 = j \,. \tag{3.5.49}$$

With the aid of (3.5.49), m_b can be eliminated from (3.5.48),

$$\begin{aligned}&\langle j_a, m_a - 1, j_b, j - m_a + 1 | j_a, j_b, j, j\rangle \\ &\quad = -\left[\frac{(j_b + j - m_a + 1)(j_b - j + m_a)}{(j_a + m_a)(j_a - m_a + 1)}\right]^{\frac{1}{2}} \langle j_a, m_a, j_b, j - m_a | j_a, j_b, j, j\rangle \,.\end{aligned} \tag{3.5.50}$$

Starting with $m_a = j_a$ in (3.5.50), all Clebsch-Gordan coefficients $\langle j_a, m_a, j_b, m_b | j_a, j_b, j, j\rangle$ can be successively calculated in terms of $\langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle$:

$$\begin{aligned}&\langle j_a, j_a - 1, j_b, j - j_a + 1 | j_a, j_b, j, j\rangle \\ &\qquad = -\left[\frac{(j_b + j - j_a + 1)(j_b - j + j_a)}{(2j_a)(1)}\right]^{\frac{1}{2}} \langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle \,.\end{aligned} \tag{3.5.51}$$

Taking $m_a = j_a - 1$ in (3.5.50),

$$\begin{aligned}&\langle j_a, j_a - 2, j_b, j - j_a + 2 | j_a, j_b, j, j\rangle \\ &= -\left[\frac{(j_b + j - j_a + 2)(j_b - j + j_a - 1)}{(2j_a - 1)(2)}\right]^{\frac{1}{2}} \langle j_a, j_a - 1, j_b, j - j_a + 1 | j_a, j_b, j, j\rangle \,.\end{aligned} \tag{3.5.52}$$

Substituting the matrix element on the left-hand side of (3.5.51) into the right-hand side of (3.5.52)

$$\begin{aligned}&\langle j_a, j_a - 2, j_b, j - j_a + 2 | j_a, j_b, j, j\rangle \\ &= \left[\frac{[(j_b + j - j_a + 1)(j_b + j - j_a + 2)][(j_b - j + j_a)(j_b - j + j_a - 1)]}{[(2j_a)(2j_a - 1)](1 \cdot 2)}\right]^{\frac{1}{2}} \\ &\qquad \times \langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle . \end{aligned} \tag{3.5.53}$$

Matrix elements have now been calculated in terms of $\langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle$ for $m_a = j_a - 1$ in (3.5.51) and for $m_a = j_a - 2$ in (3.5.53). When $m_a = j_a - n$, where n is an integer, the above equation immediately generalizes to

$$\begin{aligned}&\langle j_a, j_a - n, j_b, j - j_a + n | j_a, j_b, j, j\rangle \\ &= (-1)^n \left[\frac{(j_b + j - j_a + 1)(j_b + j - j_a + 2) \cdots (j_b + j - j_a + n)}{(2j_a)(2j_a - 1) \cdots (2j_a + 1 - n)}\right. \\ &\quad \left. \times \frac{(j_b - j + j_a)(j_b - j + j_a - 1) \cdots (j_b - j + j_a + 1 - n)}{1 \cdot 2 \cdots n}\right]^{\frac{1}{2}} \\ &\quad \times \langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle \\ &= (-1)^n \left[\frac{(j_b + j - j_a + n)!}{(j_b + j - j_a)!} \frac{(2j_a - n)!}{(2j_a)!} \frac{(j_b - j + j_a)!}{(j_b - j + j_a - n)!} \frac{1}{n!}\right]^{\frac{1}{2}} \\ &\qquad \times \langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle . \end{aligned} \tag{3.5.54}$$

Using $m_a = j_a - n$ to eliminate n from the above equation,

$$\begin{aligned}&\langle j_a, m_a, j_b, j - m_a | j_a, j_b, j, j\rangle \\ &= (-1)^{j_a - m_a} \left[\frac{(j_b + j - m_a)!(j_a + m_a)!(j_b - j + j_a)!}{(j_b + j - j_a)!(2j_a)!(j_b - j + m_a)!(j_a - m_a)!}\right]^{\frac{1}{2}} \\ &\qquad \times \langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle . \end{aligned} \tag{3.5.55}$$

The Clebsch-Gordan coefficient $\langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle$ that appears on the right hand side of (3.5.55) can be calculated from the orthogonality relation (3.5.20). Taking $m_a' = m_a$, $m_b' = m_b$, $m = j$ and noting that $m_a + m_b = j$, which implies that $m_b = j - m_a$, (3.5.20) becomes

$$1 = \sum_{m_a = -j_a}^{j_a} |\langle j_a, m_a, j_b, j - m_a | j_a, j_b, j, j\rangle|^2 . \tag{3.5.56}$$

Substituting the expression (3.5.55) for the matrix element $\langle j_a, m_a, j_b, j - m_a | j_a, j_b, j, j, \rangle$ into (3.5.56),

$$1 = \frac{(j_b - j + j_a)!}{(j_b + j - j_a)!\langle 2j_a)!} |\langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle|^2 \times \sum_{m_a=-j_a}^{j_a} \frac{(j_b + j - m_a)!(j_a + m_a)!}{(j_b - j + m_a)!(j_a - m_a)!} . \tag{3.5.57}$$

Making use of the equality,[12]

$$\sum_{m_a=-j_a}^{j_a} \frac{(j_b + j - m_a)!(j_a + m_a)!}{(j_b - j + m_a)!(j_a - m_a)!} = \frac{(j_a + j_b + j + 1)!(-j_a + j_b + j)!(j_a - j_b + j)!}{(2j + 1)!(j_a + j_b - j)!} , \tag{3.5.58}$$

the following expression is obtained for the matrix element $\langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle$;

$$\langle j_a, j_a, j_b, j - j_a | j_a, j_b, j, j\rangle = \left[\frac{(2j_a)!(2j + 1)!}{(j_a + j_b + j + 1)!(j_a - j_b + j)!} \right]^{\frac{1}{2}} , \tag{3.5.59}$$

where a choice in phase has been made when taking the square root.

Example 3.5.4 Calculate the Clebsch-Gordan coefficient

$$C^{\ell \quad 1/2 \quad \ell-1/2}_{\ell \; -1/2 \; \ell-1/2} = \langle \ell, \ell, \frac{1}{2}, -\frac{1}{2} | \ell, \frac{1}{2}, \ell - \frac{1}{2}, \ell - \frac{1}{2}\rangle$$

from (3.5.59) and verify that the result agrees with the expression given in Table 3.4 on page 177.

Solution Taking $j_a = \ell$, $j_b = \frac{1}{2}$, $j = \ell - \frac{1}{2}$, from (3.5.59)

$$\langle \ell, \ell, \frac{1}{2}, -\frac{1}{2} | \ell, \frac{1}{2}, \ell - \frac{1}{2}, \ell - \frac{1}{2}\rangle = \sqrt{\frac{(2\ell)!(2\ell)!}{(2\ell + 1)!(2\ell - 1)!}} = \sqrt{\frac{2\ell}{2\ell + 1}} ,$$

---

[12] A. R. Edmonds, *Angular Momentum in Quantum Mechanics*. Princeton University Press, Princeton, 1957.

which agrees with the Clebsch-Gordan coefficient in Table 3.4 when $j = \ell - \frac{1}{2}$, $m_b = -\frac{1}{2}$ and $m = \ell - \frac{1}{2}$.

Substituting the expression for the matrix element given in (3.5.59) into the right hand side of (3.5.55),

$$\begin{aligned}\langle j_a, m_a, j_b, j - m_a | j_a, j_b, j, j\rangle &= (-1)^{j_a - m_a} \\ &\times \left[\frac{(2j+1)!(j_b+j_a-j)!(j_b+j-m_a)!(j_a+m_a)!}{(j+j_a+j_b+1)!(j+j_a-j_b)!(j-j_a+j_b)!(j_b-j+m_a)!(j_a-m_a)!}\right]^{\frac{1}{2}} .\end{aligned} \tag{3.5.60}$$

The above equation gives all Clebsch-Gordan coefficients when $m = j$.

Example 3.5.5 Calculate the Clebsch-Gordan coefficient

$$C^{\ell}_{\ell-1} {}^{1/2}_{1/2} {}^{\ell-1/2}_{\ell-1/2} = \langle \ell, \ell-1, \frac{1}{2}, \frac{1}{2} | \ell, \frac{1}{2}, \ell - \frac{1}{2}, \ell - \frac{1}{2}\rangle$$

from (3.5.60) and verify that the result agrees with the expression given in Table 3.4.

Solution Taking $j_a = \ell, m_a = \ell - 1, j_b = \frac{1}{2}, j = \ell - \frac{1}{2}$, from (3.5.60)

$$\langle \ell, \ell-1, \frac{1}{2}, \frac{1}{2} | \ell, \frac{1}{2}, \ell - \frac{1}{2}, \ell - \frac{1}{2}\rangle = -\left[\frac{(2\ell)!(1)!(1)!(2\ell-1)!}{(2\ell+1)!(2\ell-1)!(0)!(0)!(1)!}\right]^{\frac{1}{2}} = -\frac{1}{\sqrt{2\ell+1}} ,$$

which agrees with the expression for the Clebsch-Gordan coefficient in Table 3.4 when $j = \ell - \frac{1}{2}$, $m_b = -\frac{1}{2}$ and $m = \ell - \frac{1}{2}$.

Formulas for Clebsch-Gordan coefficients when $m \neq j$ can be calculated using the lowering operator J_-. Equating (3.5.47a) and (3.5.47b) for the matrix element of J_- when $m = j$,

$$\begin{aligned}\langle j_a, m_a, j_b, m_b | j, j-1\rangle &= \sqrt{\frac{(j_a - m_a)(j_a + m_a + 1)}{2j}} \langle j_a, m_a + 1, j_b, m_b | j_a, j_b, j, j\rangle \\ &+ \sqrt{\frac{j_b - m_b)(j_b + m_b + 1)}{2j}} \langle j_a, m_a, j_b, m_b + 1 | j_a, j_b, j, j\rangle .\end{aligned} \tag{3.5.61}$$

The above equation expresses Clebsch-Gordan coefficients with $m = j - 1$ in terms of Clebsch-Gordan coefficients with $m = j$, which can be calculated from (3.5.60).

Repeating this process with $m = j-1, j-2, \cdots$, the following general formula for the Clebsch-Gordan coefficients is obtained[13]:

$$
\begin{aligned}
C^{j_a \quad j_b \quad j}_{m_a \quad m_b \quad m} &\equiv \langle j_a, m_a, j_b, m_b | j_a, j_b, j, m\rangle \\
&= \delta_{m_a+m_b,m}\left[\frac{(2j+1)(j_a+j_b-j)!(j_a-j_b+j)!(-j_a+j_b+j)!}{(j_a+j_b+j+1)!}\right]^{\frac{1}{2}} \\
&\quad \times [(j_a+m_a)!(j_a-m_a)!(j_b+m_b)!(j_b-m_b)!(j+m)!(j-m)!]^{\frac{1}{2}} \\
&\quad \times \sum_n \{(-1)^n/[n!(j_a+j_b-j-n)!(j_a-m_a-n)!(j_b+m_b-n)! \\
&\qquad \times (j-j_b+m_a+n)!(j-j_a-m_b+n)!]\} \qquad (3.5.62)
\end{aligned}
$$

The summation $n = 0, 1, 2, \cdots, N$ ranges over all integer values for which the argument of every factorial involving n is nonnegative.

Example 3.5.6 Calculate the Clebsch-Gordan coefficient

$$
C^{\ell \qquad 1/2 \quad \ell+1/2}_{m-1/2 \quad 1/2 \quad m} = \langle \ell, m-\frac{1}{2}, \frac{1}{2}, \frac{1}{2} | \ell, \frac{1}{2}, \ell+\frac{1}{2}, m\rangle
$$

from (3.5.62).

Solution Taking $j_a = \ell, m_a = m - \frac{1}{2}, j_b = \frac{1}{2}, m_b = \frac{1}{2}, j = \ell + \frac{1}{2}$, it follows that the term $(j_a + j_b - j - n)$ in the sum in (3.5.62) equals $(-n)!$. Thus the the sum is only over $n = 0$.

$$
\begin{aligned}
\langle \ell, m-\frac{1}{2}, \frac{1}{2}, \frac{1}{2} | \ell, \frac{1}{2}, \ell+\frac{1}{2}, m\rangle &= \left[\frac{(2\ell+2)(0)!(2\ell)!(1)!}{(2\ell+2)!}\right]^{\frac{1}{2}} \\
&\quad \times [(\ell+m-\frac{1}{2})!(\ell-m+\frac{1}{2})!(1)!(0)!(\ell+m+\frac{1}{2})!(\ell-m+\frac{1}{2})!]^{\frac{1}{2}} \\
&\quad \times \frac{(-1)^0}{(0)!(-0)!(\ell-m+\frac{1}{2}!(1)!(\ell+m-\frac{1}{2})!(0)!} = \sqrt{\frac{\ell+m+\frac{1}{2}}{2\ell+1}},
\end{aligned}
$$

which agrees with the expression for the Clebsch-Gordan coefficient in Table 3.4.

It is frequently more convenient to use Wigner 3-j symbols instead of Clebsch-Gordan coefficients since 3-j symbols display symmetry properties more clearly. The 3-j symbol is defined by

$$
\begin{pmatrix} j_a & j_b & j_c \\ m_a & m_b & m_c \end{pmatrix} \equiv \frac{(-1)^{j_a-j_b-m_c}}{\sqrt{2j_c+1}} \langle j_a, m_a, j_b, m_b | j_a, j_b, j_c, -m_c\rangle . \qquad (3.5.63)
$$

---

[13] *Ibid*, pp. 44, 45.

Its symmetry properties are given by

$$\begin{pmatrix} j_a & j_b & j_c \\ m_a & m_b & m_c \end{pmatrix} = \begin{pmatrix} j_b & j_c & j_a \\ m_b & m_c & m_a \end{pmatrix} = \begin{pmatrix} j_c & j_a & j_b \\ m_c & m_a & m_b \end{pmatrix}, \tag{3.5.64a}$$

$$(-1)^{j_a+j_b+j_c} \begin{pmatrix} j_a & j_b & j_c \\ m_a & m_b & m_c \end{pmatrix}$$
$$= \begin{pmatrix} j_b & j_a & j_c \\ m_b & m_a & m_c \end{pmatrix} = \begin{pmatrix} j_a & j_c & j_b \\ m_a & m_c & m_b \end{pmatrix} = \begin{pmatrix} j_c & j_b & j_a \\ m_c & m_b & m_a \end{pmatrix}, \tag{3.5.64b}$$

and

$$\begin{pmatrix} j_a & j_b & j_c \\ m_a & m_b & m_c \end{pmatrix} = (-1)^{j_a+j_b+j_c} \begin{pmatrix} j_a & j_b & j_c \\ -m_a & -m_b & -m_c \end{pmatrix}. \tag{3.5.64c}$$

3.6 Tensor Operators and the Wigner-Eckart Theorem

Angular momentum was considered in some detail in Chap. 2, and Clebsch-Gordan coefficients were discussed in Sect. 3.5. Here a more general class of operators, called *tensor operators*,[14] are defined by their relationship with the angular-momentum operators. The simplest example of such an operator is a *scalar operator* defined to be any operator S satisfying

$$[J_i, S] = 0. \tag{3.6.1}$$

Another example is any set of three operators $V_i (i = 1, 2, 3)$ that satisfy

$$[J_i, V_j] = i\hbar\epsilon_{ijk} V_k, \quad i, j, k = 1, 2, 3. \tag{3.6.2}$$

Such a set of operators is called a *vector operator* or *regular tensor operator*. Note that the angular momentum itself is a vector operator. Scalar and vector operators are tensor operators of rank 0 and 1, respectively.

Rather than using the "Cartesian" components V_i, it is more convenient to use "spherical" components V_0 and $V_{\pm1}$ defined by

$$V_0 = V_3, \quad V_{\pm1} = \frac{\mp1}{\sqrt{2}}(V_1 \pm i V_2). \tag{3.6.3}$$

---

[14] Here only "irreducible tensor operators" are considered.

When written in terms of spherical components, (3.6.2) becomes

$$[J_\kappa, V_\kappa] = 0\,, \quad \kappa = -1, 0, +1\,, \tag{3.6.4a}$$

$$[J_\pm, V_\mp1] = \sqrt{2}\hbar V_0\,, \quad [J_\pm, V_0] = \sqrt{2}\hbar V_{\pm 1}\,, \quad [J_0, V_\kappa] = \kappa\hbar V_\kappa\,. \tag{3.6.4b}$$

In general, when written in spherical components, a *tensor operator* $T^{(j)}_\kappa$ ($\kappa = -j, -j+1, \ldots, +j$) of rank j is a set of $2j+1$ operators[15] that satisfy

$$[J_0\,,\ T^{(j)}_\kappa] = \kappa\hbar\, T^{(j)}_\kappa\,, \tag{3.6.5a}$$

$$[J_\pm\,, T^{(j)}_\kappa] = \sqrt{j(j+1) - \kappa(\kappa\pm 1)}\,\hbar\, T^{(j)}_{\kappa\pm 1}\,. \tag{3.6.5b}$$

As can be checked (Problem 3.32), using the Clebsch-Gordan coefficients $C^{j\ \ 1\ \ \ j}_{\kappa\ \mu\ \kappa+\mu}$ listed in Table 3.5 on page 177, the above equation can be written in the following compact form:

$$[J_\mu,\ T^{(j)}_\kappa] = \sqrt{j(j+1)}\ C^{j\ \ 1\ \ \ j}_{\kappa\ \mu\ \kappa+\mu}\,\hbar\, T^{(j)}_{\kappa+\mu} \tag{3.6.6}$$

The matrix elements of tensor operators have an important property which is expressed by the *Wigner-Eckart theorem*: Let $T^{(J)}_\kappa$ be a tensor operator. The matrix element of $T^{(J)}_\kappa$ between angular momentum eigenstates may be written as

$$\langle j', m'|T^{(J)}_\kappa|j, m\rangle = C^{j\ \ J\ \ j'}_{m\ \kappa\ m'}\,\langle j'\|T^{(J)}\|j\rangle\,, \tag{3.6.7}$$

where $C^{j\ \ J\ \ j'}_{m\ \kappa\ m'}$ is a Clebsch-Gordan coefficient. The quantity $\langle j'\|T^{(J)}\|j\rangle$ defined by (3.6.7) is called the reduced matrix element of the tensor operator and depends on j', j, J, and the nature of the tensor operator $T^{(J)}_\kappa$, but it does not depend on m, m', or κ.[16] This theorem will not be proven here since the proof is pure mathematics and does not give any additional insight into the physics.

A particular consequence of the Wigner-Eckart theorem (3.6.7) is

$$\langle j', m'|T^{(J)}_\kappa|j, m\rangle = 0 \ \text{ if } m' \neq \kappa + m \ \text{ or } j' \neq J+j,\ J+j-1,\ \ldots,\ |J-j|\,, \tag{3.6.8}$$

---

[15] In labeling the components of vector operators, spherical components are denoted by Greek letters and Cartesian components by Latin letters. This notation is not used for the general tensor operator.

[16] In some textbooks, a constant factor or function of j appears explicitly in the Wigner-Eckart theorem. In the notation used here these factors have been absorbed into the reduced matrix element.

which follows immediately from the properties of the Clebsch-Gordan coefficients (3.5.25). For a scalar operator S, (3.6.7) becomes

$$\langle j', m'|S|j, m\rangle = \delta_{m'm}\delta_{j'j}\langle j\|S\|j\rangle \,. \tag{3.6.9}$$

The above equation reveals that S cannot change the angular-momentum quantum number, an expected result.

Example 3.6.1 Calculate the reduced matrix element $\langle j'\|T^{(J)}\|j\rangle$ for $T^{(J)}_{\kappa=0} = J_0$.

Solution For the angular-momentum operators themselves, (3.6.7) gives

$$\langle j', m'|J_0|j, m\rangle = C^{j\ \ 1\ \ j'}_{m\ 0\ m'}\langle j'\|J\|j\rangle \,. \tag{3.6.10}$$

The left-hand side of (3.6.7) is evaluated using (2.3.43),

$$\langle j', m'|J_0|j, m\rangle = \hbar\, m\, \delta_{j',j}\, \delta_{m',m} \,, \tag{3.6.11a}$$

and the right-hand side of (3.6.7) is simplified using the explicit expressions for the Clebsch-Gordan coefficient given in Table 3.5 on page 177:

$$\langle j', m'|J_0|j, m\rangle = \frac{m}{\sqrt{j(j+1)}}\langle j'\|J\|j\rangle\, \delta_{m',m} \,. \tag{3.6.11b}$$

Equating (3.6.11a) and (3.6.11b) ,

$$\langle j'\|J\|j\rangle = \delta_{j'j}\hbar\sqrt{j(j+1)}\,. \tag{3.6.12}$$

Identical expressions for $\langle j'\|J\|j\rangle$ are, of course, obtained when $T^{(J)}_{\kappa=\pm1} = J_{\pm1}$ (Problem 3.33).

If, in addition to j and m, there are other quantum numbers, say $\eta = a_1\,, a_2\,, \ldots..,\, a_N$, then in general the reduced matrix element will also depend on the quantum numbers η. That is,

$$\langle \eta', j', m'|T^{(J)}_\kappa|\eta, j, m\rangle = C^{j\ \ J\ \ j'}_{m\ \kappa\ m'}\langle \eta', j'\|T^{(J)}\|\eta, j\rangle \,. \tag{3.6.13}$$

The additional observables $\eta^{\text{op}} = A_1,\ A_2,\ \ldots,\ A_N$ whose respective eigenvalues are the quantum numbers $\eta = a_1,\ a_2,\ \ldots,\ a_N$ must have the property that $[\eta^{\text{op}}, J_i] = 0$ or $[\eta^{\text{op}}, \text{SO}(3)_{J_i}] = 0$. It should be emphasized that the Wigner-Eckart theorem is both a theorem and a definition. It is a theorem because (3.6.7) reveals that the matrix element of $T^{(J)}_\kappa$ can be factored so that the dependence on m, m', and κ is contained entirely in the Clebsch-Gordan coefficient. It is a definition because (3.6.7) defines the reduced matrix elements.

The Wigner-Eckart theorem has become one of the more important tools for the understanding of physics. Equation (3.6.7) is the Wigner-Eckart theorem for the rotation group, which is connected with the algebra of angular momenta J_i as an algebra of observables. Many physical systems have an (enveloping) algebra of a group as a subalgebra of the algebra of observables, and they have observables that are tensor operators with respect to this group. For those observables, a generalization of the the Wigner-Eckart theorem (3.6.7) is valid.

The Wigner-Eckart theorem expresses the matrix elements of the tensor operators, which are directly connected with the numbers measured in an experiment, in terms of Clebsch-Gordan coefficients and reduced matrix elements. The Clebsch-Gordan coefficients are known mathematical quantities and are calculated from the properties of the group. The reduced matrix elements are physical parameters with values that have to be obtained from the experimental data. For an observable such as Q_j that obeys the additional relation (2.4.1c), the reduced matrix elements can be reduced to a still smaller number of parameters. The importance of the Wigner-Eckart theorem is that it allows the large number of experimentally observable matrix elements to be expressed in terms of the reduced matrix elements, of a much smaller number of more fundamental quantities. Often all that is known about an observable is that it is a tensor operator. The Wigner-Eckart theorem then is the only tool at hand.

If $|j, m\rangle$ is a basis in the space in which V_κ is a general vector operator, then $V_\kappa|j, m\rangle$ can be expanded with respect to $|j, m\rangle$:

$$V_\kappa|j, m\rangle = \sum_{j', m'} |j', m'\rangle\langle j', m'|V_\kappa|j, m\rangle \tag{3.6.14}$$

According to the Wigner-Eckart theorem (3.6.7),

$$V_\kappa|j, m\rangle = \sum_{j', m'} |j', m'\rangle C^{j\ \ 1\ \ j'}_{m\ \kappa\ m'}\langle j'\|V\|j\rangle\,. \tag{3.6.15}$$

If the c. s. c. o. contains, in addition to $\mathbf{J}^2$ and J_3, N other operators $A_1, \ldots, A_N$ with respective spectra $\eta = (a_1,\ a_2,\ \ldots,\ a_N)$, then the basis is $|\eta, j, m\rangle$. Instead of (3.6.15) one has, according to (3.6.13),

$$V_\kappa|\eta, j, m\rangle = \sum_{\eta' j' m'} |\eta', j', m'\rangle C^{j\ \ 1\ \ j'}_{m\ \kappa\ m'}\langle \eta', j'\|V\|\eta, j\rangle\,. \tag{3.6.16}$$

In the remainder of this section the additional quantum numbers η will be suppressed, but it should be understood that whenever additional quantum numbers η are needed, the reduced matrix elements will depend on these quantum numbers and a summation over η' is implied.

According to (3.6.8), the only non-vanishing terms in (3.6.15) are those for which $j' = j+1\,,\ j,\ j-1$ and $m' = m+\kappa$:

$$V_\kappa|j,m\rangle = |j-1,m+\kappa\rangle C^{j\ \ 1\ \ j-1}_{m\ \kappa\ m+\kappa}\langle j-1\|V\|j\rangle + |j,m+\kappa\rangle C^{j\ \ 1\ \ j}_{m\ \kappa\ m+\kappa}\langle j\|V\|j\rangle + |j+1,m+\kappa\rangle C^{j\ \ 1\ \ j+1}_{m\ \kappa\ m+\kappa}\langle j+1\|V\|j\rangle\,. \tag{3.6.17}$$

This is the most general form possible for the action of a vector operator. According to (3.6.17) any vector operator can be completely specified by three reduced matrix elements that in general may depend on η, η'. Actually, as shall be shown below, only two quantities are needed to determine a vector operator.

Using the Clebsch-Gordan coefficients in Table 3.5 on page 177, (3.6.17) can be written explicitly. For the 0-component,

$$\begin{aligned} V_0|j,m\rangle &= |j-1,m\rangle\left(-\sqrt{\frac{(j-m)(j+m)}{j(2j+1)}}\right)\langle j-1\|V\|j\rangle \\ &\quad + |j,m\rangle\left(\frac{m}{\sqrt{j(j+1)}}\right)\langle j\|V\|j\rangle \\ &\quad + |j+1,m\rangle\left(\sqrt{\frac{(j-m+1)(j+m+1)}{(2j+1)(j+1)}}\right)\langle j+1\|V\|j\rangle\,, \end{aligned} \tag{3.6.18}$$

Defining c_j, a_j, and d_j,

$$c_j = -\frac{\langle j-1\|V\|j\rangle}{\sqrt{j(2j+1)}}\,, \tag{3.6.19a}$$

$$a_j = -\frac{\langle j\|V\|j\rangle}{\sqrt{j(j+1)}}\,, \tag{3.6.19b}$$

$$d_j = -\frac{\langle j+1\|V\|j\rangle}{\sqrt{(2j+1)(j+1)}}\,, \tag{3.6.19c}$$

(3.6.18) can be rewritten as

$$V_0|j,m\rangle = \sqrt{j^2-m^2}c_j|j-1,m\rangle - ma_j|j,m\rangle - \sqrt{(j+1)^2-m^2}d_j|j+1,m\rangle\,. \tag{3.6.20}$$

The equations for $V_{\pm1}|j,m\rangle$ can similarly be rewritten as

$$\begin{aligned} V_{+1}|j,m\rangle &= -|j-1,m+1\rangle\sqrt{(j-m-1)(j-m)/2}\,c_j \\ &\quad + |j,m+1\rangle\sqrt{(j+m+1)(j-m)/2}\,a_j \\ &\quad - |j+1,m+1\rangle\sqrt{(j+m+1)(j+m+2)/2}\,d_j\,, \end{aligned} \tag{3.6.21}$$

and

$$\begin{aligned} V_{-1}|j,m\rangle = -|j-1,m-1\rangle\sqrt{(j+m)(j+m-1)/2}\,c_j \\ -|j,m-1\rangle\sqrt{(j-m+1)(j+m)/2}\,a_j \\ -|j+1,m-1\rangle\sqrt{(j-m+1)(j-m+2)/2}\,d_j\,. \end{aligned} \tag{3.6.22}$$

Equations (3.6.20)–(3.6.22) reveal that a vector operator is completely determined by the three functions c_j, a_j, and d_j that depend on the discrete parameter j and may in general depend on η' and η. In the mathematical note that follows, it will be shown that it is possible to choose a different angular momentum basis in such a way that only two functions are required.

3.6.1 Mathematical Note

A new angular momentum basis

$$|h_m^j\rangle = \xi(j)|j,m\rangle\,, \tag{3.6.23}$$

is defined where $\xi(j)$ is a complex number. The action of the angular momentum operators on the basis $|h_m^i\rangle$ is identical to their action (2.3.43) on the basis $|j,m\rangle$. In general, the $|h_m^j\rangle$ are not normalized unless

$$|\xi(j)| = 1\,. \tag{3.6.24}$$

Multiplying (3.6.20) by $\xi(j)$, multiplying and dividing the first term on the right-hand side by $\xi(j-1)$, and multiplying and dividing the third term on the right-hand side by $\xi(j+1)$,

$$\begin{aligned} V_0|h_m^j\rangle = \sqrt{j^2-m^2}\,c_j\frac{\xi(j)}{\xi(j-1)}|h_m^{j-1}\rangle - ma_j|h_m^j\rangle \\ -\sqrt{(j+1)^2-m^2}\,d_j\frac{\xi(j)}{\xi(j+1)}|h_m^{j+1}\rangle\,. \end{aligned} \tag{3.6.25}$$

Defining

$$A_j \equiv a_j\,, \tag{3.6.26a}$$

$$C_j \equiv c_j\frac{\xi(j)}{\xi(j-1)}\,, \tag{3.6.26b}$$

$$D_j \equiv d_j\frac{\xi(j)}{\xi(j+1)}\,, \tag{3.6.26c}$$

the complex number $\xi(j)$ is chosen so that $C_j = D_{j-1}$:

$$c_j \frac{\xi(j)}{\xi(j-1)} = \frac{\xi(j-1)}{\xi(j)} d_{j-1}\,, \tag{3.6.27a}$$

or, equivalently,

$$\xi^2(j) = \xi^2(j-1)\frac{d_{j-1}}{c_j}\,. \tag{3.6.27b}$$

Suppose j_0 is the smallest value of j in the space so that

$$\langle j_0 - 1 \| V \| j_0 \rangle = 0\,. \tag{3.6.28}$$

Then from (3.6.19) it follows that $c_j = 0$ and $d_{j-1} = 0$ for $j < j_0$. By direct substitution it is easy to see that (3.6.27b) is satisfied if

$$\xi(j) = \sqrt{\xi^2(j_0)\frac{d_{j_0}}{c_{j_0+1}}\frac{d_{j_0+1}}{c_{j_0+2}}\cdots\frac{d_{j-1}}{c_j}}\,. \tag{3.6.29}$$

Equation (3.6.25) can now be rewritten as

$$V_0|h^j_m\rangle = \sqrt{j^2-m^2}C_j|h^{j-1}_m\rangle - mA_j|h^j_m\rangle - \sqrt{(j+1)^2-m^2}C_{j+1}|h^{j+1}_m\rangle\,. \tag{3.6.30}$$

Using (3.6.27a) and (3.6.29) to express d_{j-1} in terms of c_j, the action of every vector operator acting on an angular momentum basis system can be written in the form (3.6.20) to (3.6.22) where the a_j and c_j are functions of j. It should be noted, however, that in general it is possible have $C_j = D_{j-1}$ for only one vector operator V_κ at a time. If two different vector operators are involved in one problem and the basis has been chosen so that for the reduced matrix element of one of them $C_j = D_{j-1}$, then in this basis, the other or any additional vector operator is expressed in terms of three independent reduced matrix elements.

So far the only condition placed on V_κ is that it be a vector operator. As a consequence the reduced matrix elements c_j and a_j are arbitrary functions of j that cannot be calculated: they can only be determined phenomenologically from the experimental value of one particular matrix element of one component of V_κ. Specifically, c_j can be obtained from $\langle j-1, j|V_0|j, j\rangle$ and a_j from $\langle j, j|V_0|j, j\rangle$. Using the Clebsch-Gordan coefficients all other matrix elements $\langle j, m|V|j, m\rangle$ for $m = -j, -j+1, \ldots, j$ and $\kappa = 0, +1, -1$ can then be calculated from these two experimental values.

If V_κ is required to obey additional conditions, more specific information about the c_j and a_j is obtained. For example, as was shown in Example 3.6.1 on page 185, the reduced matrix elements were completely determined when the vector operators were the angular momentum operators.

Often the vector operators have a specified hermiticity property. They are either hermitian, defined by

$$V_0^+ = V_0\,, \quad V_{\pm 1}^+ = -V_{\mp 1}\,, \tag{3.6.31}$$

or they are skew-hermitian, defined by

$$V_0^+ = -V_0\,, \quad V_{\pm 1}^+ = +V_{\mp 1}\,. \tag{3.6.32}$$

It is easy to see (Problem 3.36) that for Hermetian V_κ the functions c_j are purely imaginary, $(c_j)^* = -c_j$, and the functions a_j are real, $(a_j)^* = a_j$. For skew-Hermetian V_κ one has $(c_j)^* = c_j$ and $(a_j)^* = -a_j$.

3.7 Summary

If a quantum particle has no structure and is truly a point object, it is fundamental and is called an elementary particle. For particles with a finite spatial extent, whether they are thought of as being fundamental or composite depends upon the energy used to study them.

When two mass points interact via a potential that depends only upon the relative positions of the mass points, the original two-body Hamiltonian $\hat{H}$,

$$\hat{H} = \frac{\mathbf{P}_1^2}{2m_1} + \frac{\mathbf{P}_2^2}{2m_2} + U(\mathbf{Q}_2 - \mathbf{Q}_1)\,.$$

can be rewritten as a sum of two, one-body Hamiltonians by rewriting the original Hamiltonian in terms of the relative position operator $\mathbf{Q}$, the relative momentum operator $\mathbf{P}$ and the center-of-mass momentum operator $\mathbf{P}_{\mathrm{CM}}$ where

$$\begin{aligned}
\mathbf{Q} &= \mathbf{Q}_2 - \mathbf{Q}_1\,,\\
\mathbf{P} &= \frac{m_1\mathbf{P}_2 - m_2\mathbf{P}_1}{M}\,,\\
\mathbf{P}_{\mathrm{CM}} &= \mathbf{P}_1 + \mathbf{P}_2\,.
\end{aligned}$$

In terms of these new operators, the original Hamiltonian $\hat{H}$ takes the form

$$H = \frac{\mathbf{P}_{\mathrm{CM}}^2}{2M} + \frac{\mathbf{P}^2}{2\mu} + U(\mathbf{Q})\,,$$

The first term is the Hamiltonian of a free particle with mass M and the sum of the second and third terms is the Hamiltonian of a particle with mass μ in an external potential $U(\mathbf{Q})$. By changing to the center-of-mass and relative variables, the two-body Hamiltonian is reduced to two, one-body Hamiltonians, one of which is trivial to solve.

The basic postulate for combining quantum physical systems is as follows: Let one physical system be described by a set of operators $\{A_\Phi\}$ in the space Φ, and a second physical system be described by a set of operators $\{B_\Psi\}$ in the space Ψ. The direct-product space $\Phi \otimes \Psi$ is the space of physical states of the combination of the two systems. The observables C for the combined system are operators in the direct-product space and are of the form

$$C = \sum a_i A^i_\Phi \otimes B^i_\Psi \,, \quad a_i \in \mathbb{C} \,.$$

Observables in the first system alone are given by $A_\Phi \otimes 1$, and observables in the second system alone are given by $1 \otimes B_\Psi$.

A set of commuting, hermitian operators whose eigenvalues completely specify a (generalized) basis for a system is called a complete set of commuting operators or a complete set of commuting observables for the system. When studying a quantum physical system, the two most important questions are as follows: (1) What is a complete set of commuting operators for the system? (2) What additional operators must be adjoined to the complete set of commuting operators and what are the algebraic relations among the various operators? For a quantum physical system with a classical analogue, the number of operators in a complete set of commuting observables equals the number of degrees of freedom in the corresponding classical system.

Understanding of a complex physical system can often be achieved by thinking of it as a combination of two or more elementary physical systems. In quantum mechanics, complex physical system can be understood atomistically (what the complex system is made of) or it in terms of collective motions (what motions the complex system can perform).

A diatomic molecule, which is an example of a vibrating rotator when observed at sufficiently high energies, is an instructive example for the combination of two quantum physical systems. When a diatomic molecule is studied with energies $\lesssim 10^{-2}$ eV, only rotational states are excited and the molecule can be described as a rotator with a Hamiltonian

$$H_{\text{rotator}} = \frac{\mathbf{L}^2}{2I} \,.$$

When the energy is increased to about 0.1 eV, the molecule begins vibrating with the result that new energy levels are observed, each of which contains an entire rotational band. The diatomic molecule can now be described as a vibrating rotator using a direct-product space. As a first approximation the Hamiltonian H is

$$H = H_{\text{oscillator}} \otimes I + I \otimes H_{\text{rotator}} \,,$$

where

$$H_{\text{oscillator}} = \frac{\mathbf{P}^2}{2\mu} + \frac{1}{2}kQ^2 .$$

When the diatomic molecule is in a state of higher vibrational energy, it oscillates with a larger amplitude and consequently has a larger moment of inertia, implying that the energy associated with oscillations affects the energy associated with rotation. A simple term H_{int} representing interaction between vibrational and rotational degrees of freedom is

$$H_{\text{int}} = gH_{\text{oscillator}} \otimes H_{\text{rotator}} ,$$

where g is a constant. The Hamiltonian

$$H = H_{\text{oscillator}} \otimes I + I \otimes H_{\text{rotator}} + gH_{\text{oscillator}} \otimes H_{\text{rotator}} .$$

describes the energy spectra of various diatomic molecules quite well. However, agreement with empirical data can be further enhanced by including anharmonic effects of the oscillator and the influence of centripetal forces.

When two physical systems with respective angular momenta j_a and j_b are combined, the possible values of total angular momentum j are

$$j = j_a + j_b,\ j_a + j_b - 1,\ j_a + j_b - 2,\ \ldots,\ |j_a - j_b| .$$

The combined rotational motion is described by the direct-product space

$$\mathfrak{R} = \mathfrak{R}^{j_a} \otimes \mathfrak{R}^{j_b} ,$$

that is spanned by the set of direct-product vectors

$$|j_a, m_a\rangle \otimes |j_b, m_b\rangle .$$

Linear combinations of the above direct-product vectors yield new vectors that are eigenstates with total angular momentum j and 3-component of total angular momentum m,

$$|j_a, j_b, j, m\rangle = \sum_{m_a, m_b} C^{j_a\ j_b\ j}_{m_a\ m_b\ m} |j_a, m_a\rangle \otimes |j_b, m_b\rangle .$$

The transition coefficients in the above formula are called Clebsch-Gordan or Wigner coefficients.

A tensor operator $T_\kappa^{(j)}$ ($\kappa = -j, -j+1, \ldots, +j$) of rank j is a set of $2j+1$ operators that satisfy

$$[J_\mu,\, T_\kappa^{(j)}] = \sqrt{j(j+1)}\, C^{j\ \ 1\ \ \ \ j}_{\kappa\ \ \mu\ \ \kappa+\mu}\, \hbar\, T^{(j)}_{\kappa+\mu}\,. \tag{3.7.1}$$

The Wigner-Eckart theorem states that the matrix element of $T_\kappa^{(J)}$ between angular momentum eigenstates may be written as

$$\langle j', m' | T_\kappa^{(J)} | j, m\rangle = C^{j\ \ J\ \ j'}_{m\ \ \kappa\ \ m'} \langle j' \| T^{(J)} \| j\rangle\,. \tag{3.7.2}$$

The dependence of the matrix elements on m, m', and κ is contained entirely in the known Clebsch-Gordan coefficients. The quantity $\langle j' \| T^{(J)} \| j\rangle$ defined by (3.7.2) is called the reduced matrix element of the tensor operator; often can only be determined experimentally; and depends only on the nature of the tensor operator $T_\kappa^{(J)}$, j', j, and J.

Problems

For Sect. 3.2

3.1 Derive (3.2.36) using the Heisenberg commutation relations (3.2.35).

3.2 Verify that when the Hamiltonian (3.2.24) is rewritten in terms of $\mathbf{P}_{\mathrm{CM}}$, $\mathbf{P}$ and $\mathbf{Q}$ it is given by (3.2.37).

3.3 Show that

$$[L_i, \mathbf{P}^2] = 0\,,$$

where the L_i are defined by (2.2.5).

3.4 Show that

$$[L_i, \mathbf{Q}^2] = 0\,,$$

where the L_i are defined by (2.2.5).

3.5 Consider a composite system with a Hamiltonian

$$H = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m} + \frac{1}{2}k\left(Q_1^2 + Q_2^2\right) + \lambda Q_1 Q_2\,.$$

Classically this system corresponds to two identical harmonic oscillators in one-dimensional space coupled by the interaction $\lambda Q_1 Q_2$.

(a) Determine new position operators Q_A and Q_B that are linear combinations of Q_1 and Q_2 and new momentum operators P_A and P_B that are linear combinations of P_1 and P_2 such that the Hamiltonian separates and

$$[P_A, Q_A] = \frac{\hbar}{i} 1 \qquad [P_B, Q_B] = \frac{\hbar}{i} 1$$

$$[P_A, Q_B] = 0 \qquad [P_B, Q_A] = 0$$

(b) What is H in terms of these new operators?

3.6 Express the time-independent, two-body Schrödinger equation (3.2.54) in terms of relative coordinates $x = x_1 - x_2$ and center-of-mass coordinates $X_{\mathrm{CM}} = (m_1 x_1 + m_2 x_2)/(m_1 + m_2)$.

3.7

(a) Using a procedure similar to that which lead to (3.2.54), derive the Schrödinger equation in two-dimensional space that describes two particles with respective charges of e and $-e$ interacting via the Coulomb force.

(b) Rewrite the Schrödinger equation in terms of the polar variables r and ϕ where $x = r\cos\phi$ and $y = r\sin\phi$ to derive the energy eigenvalue equation

$$-\frac{\hbar^2}{2\mu}\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2}\right]\langle r, \phi | E, m\rangle = E\langle r, \phi | E, m\rangle$$

that appears in Example 3.2.2 on page 134.

3.8 Beginning with the equation

$$L_3 | E, m\rangle = \hbar m | E, m\rangle \ ,$$

where $L_3 \equiv L_x$ is given by (2.2.3a), derive the partial differential equation

$$\frac{\hbar}{i}\frac{\partial}{\partial \phi}\langle r, \phi | E, m\rangle = \hbar m \langle r, \phi | E, m\rangle \ ,$$

where $x = r\cos\phi$ and $y = r\sin\phi$.

3.9 Consider the transformation from spherical coordinates to rectangular (Cartesian) coordinates:

$$x = r\sin\theta\cos\phi \ , \quad y = r\sin\theta\sin\phi \ , \quad z = r\cos\theta \ .$$

(a) Show that the transformation from rectangular coordinates to spherical coordinates is

$$r = \sqrt{x^2 + y^2 + z^2}\,, \qquad \phi = \tan^{-1}\frac{y}{x}\,, \qquad \theta = \sin^{-1}\frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}}\,.$$

(b) Using the chain rule for differentiation,

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} + \frac{\partial \phi}{\partial x}\frac{\partial}{\partial \phi}$$

etc., show that

$$\begin{aligned}
\frac{\partial}{\partial x} &= \sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\phi\frac{\partial}{\partial \theta} - \frac{1}{r}\frac{\sin\phi}{\sin\theta}\frac{\partial}{\partial \phi}\,,\\
\frac{\partial}{\partial y} &= \sin\theta\sin\theta\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\phi\frac{\partial}{\partial \theta} - \frac{1}{r}\frac{\cos\phi}{\sin\theta}\frac{\partial}{\partial \phi}\,,\\
\frac{\partial}{\partial z} &= \cos\theta\frac{\partial}{\partial r} - \frac{1}{r}\frac{\partial}{\partial \theta}\,.
\end{aligned}$$

(c) Show that

$$\begin{aligned}
\langle\theta,\phi|\mathbf{L}^2|l,m\rangle &= \langle\theta,\phi|L_1^2 + L_2^2 + L_3^2|l,m\rangle\\
&= -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\langle\theta,\phi|l,m\rangle\,.
\end{aligned}$$

For Sect. 3.3

3.10 Let T, $\tilde{T}$, and T' be tensors in the direct-product space $\Phi_M \otimes \Psi_N$. Show that the scalar product defined in (3.3.5) possesses the following properties required of a scalar product in a linear, scalar-product space:

(a) $(T, T) \geq 0$
(b) $(T, T) = 0$ iff $T = 0$
(c) $(T, \tilde{T}) = (\tilde{T}, T)^*$
(d) $a(T, \tilde{T}) = (T, a\tilde{T}) = (a^*T, \tilde{T}), \qquad a \in \mathbb{C}$
(e) $(T' + \tilde{T}, T) = (T', T) + (\tilde{T}, T)$

3.11 Let T and $\tilde{T}$ be tensors in the scalar-product space $\Phi_M \otimes \Psi_N$. Also let A_Φ, A'_Φ, B_Ψ, and B'_Ψ be linear operators in the indicated linear, scalar-product spaces. In addition, let

$$C = A_\Phi \otimes B_\Psi\,, \qquad C' = A'_\Phi \otimes B'_\Psi$$

be operators in the direct-product space. Show that the operators C and C' possess the following properties required of a linear operator in the direct-product space:

(a) $C(T+\tilde{T}) = CT + C\tilde{T}$
(b) $C(aT) = a(CT), \qquad a \in \mathbb{C}$
(c) $(C+C')T = CT + C'T$
(d) $(aC)T = a(CT), \qquad a \in \mathbb{C}$
(e) $(C'C)T = C'(CT)$

3.12 Verify Eq. (3.3.9).

3.13 Let A_Φ be a linear operator in the space Φ spanned by the orthonormal vectors ϕ_1 and ϕ_2. Similarly, let B_Ψ be a linear operator in the space Ψ spanned by the orthonormal vectors ψ_1 and ψ_2. The matrix elements of A_Φ and B_Ψ are given, respectively, by

$$\langle\phi_i|A_\Phi|\phi_j\rangle = a_{ij}\,, \qquad \langle\psi_i|A_\Phi|\psi_j\rangle = b_{ij}$$

Write the direct product basis as the column vector

$$\chi = \begin{pmatrix} \phi_1 \otimes \psi_1 \\ \phi_1 \otimes \psi_2 \\ \phi_2 \otimes \psi_1 \\ \phi_2 \otimes \psi_2 \end{pmatrix}$$

and determine the 4×4 matrix $M(a_{ij}, b_{ij})$ such that

$$A_\Phi \otimes B_\Psi\, \chi = M(a_{ij}, b_{ij})\, \chi\,.$$

Hint:

$$A\phi_i = \sum_{j=1}^{2} \phi_j a_{ji}\,, \qquad B\psi_i = \sum_{j=1}^{2} \psi_j b_{ji}\,,$$

3.14 Since a diatomic molecule can be viewed as a bound state of two mass points, from the discussion in Sect. 3.3 it is expected that six eigenvalues would be required to completely specify the state of a diatomic molecule. However, in Chap. 2, Sect. 2.4, a rotating diatomic molecule was described by the basis vectors $|\ell, m\rangle$, which are characterized by only two quantum numbers. Explain why the four additional quantum numbers were not required.

For Sect. 3.4

3.15 Calculate the average value of the moment of inertia $I = \mu(r_0 + Q)^2$ for a quantum harmonic oscillator with mass μ and energy $E_n = \hbar\omega(n + 1/2)$.

3.16 For the HCl molecule, what is the value of the constant g that appears in (3.4.21) and (3.4.23)?

For Sect. 3.5

3.17 If $J_i = J_{ai} + J_{bi}$, verify the J_i satisfy the algebra of angular momentum provided that $[J_{ai}, J_{bj}]= 0$ and that the J_{ai} and J_{bi} independently satisfy the algebra of angular momentum.

3.18 Derive the expression for $\mathbf{J}^2$ given in (3.5.13a).

3.19 Show that the eigenvector $|j_a, m_a = j_a\rangle \otimes |j_b, m_b = j_b\rangle$ is an eigenvector of $\mathbf{J}^2$. What is the value of j for the eigenvector? (Hint: $J_{a+}|j_a, j_a\rangle = J_{b+}|j_b, j_b\rangle = 0$. Why?)

3.20

(a) Write the direct product $\Re^3 \otimes \Re^2$ as a direct sum as is done in (3.5.24).
(b) What is the dimension of the space $\Re^j$? That is, how many linearly independent basis vectors are required to span the space $\Re^j$?
(c) What is the dimension of the direct-product space $\Re^3 \otimes \Re^2$?
(d) Sum the dimensions of each space on the right-hand side of (a) and verify that the sum equals the answer to (c).

3.21 The number of linearly independent eigenvectors $|j_a, j_b, j, m, \rangle$ is

$$\sum_{j_a-j_b}^{j_a+j_b} (2j + 1)\,,$$

where $j_a > j_b$. Show that this sum equals $(2j_a + 1)(2j_b + 1)$. Hint:

$$\sum_{n=1}^{N} n = \frac{1}{2}N(N + 1)\,.$$

3.22 Using the normalization condition $\langle j, m|j', m'\rangle = \delta_{m,m'}\delta_{j,j'}$, show that the vectors $|1/2, 1/2, 0, 0\rangle$ and $|1/2, 1/2, 1, 0\rangle$ given in (3.5.34) and (3.5.35), respectively, are orthogonal and that each is normalized to unity.

3.23 Derive equation (3.5.35) by applying J_+ to $|1/2, 1/2, 1, -1\rangle$ given in (3.5.31b).

3.24

(a) Take the scalar product of (3.5.18) with $|j_a, j_b, j', m'\rangle$ and verify that

$$\delta_{j,j'}\delta_{m,m'} = \sum_{m_a,m_b} C^{j_a\ j_b\ j'}_{m_a\ m_b\ m'}\, C^{j_a\ j_b\ j}_{m_a\ m_b\ m}\,.$$

(b) Using the explicit expressions for Clebsch-Gordan coefficients in (3.5.41), verify the above equation is satisfied for $j_a = j_b = 1/2$, $j = j' = 1$ and $m = m' = 0$.

3.25

(a) Using Table 3.4 on page 177 write expressions for the following Clebsch-Gordon coefficients:

$$C^{j\quad\ \ 1/2\quad j-1/2}_{m-1/2\ \ 1/2\quad\ m}\,,\qquad C^{j\quad\ \ \ 1/2\quad j-1/2}_{m+1/2\ \ -1/2\quad\ m}\,,$$

$$C^{j\quad\ \ 1/2\quad j+1/2}_{m-1/2\ \ 1/2\quad\ m}\,,\qquad C^{j\quad\ \ \ 1/2\quad j+1/2}_{m+1/2\ \ -1/2\quad\ m}\,.$$

(b) If j = 1/2, calculate the non-zero Clebsch-Gordan coefficients and verify that they agree with the expressions in (3.5.41).

3.26 Calculate the Clebsch-Gordan coefficient $C^{j_a\ \ 1\ \ j_a-1}_{j_a\ -1\ j_a-1}$ using (3.5.59) and verify that the answer agrees with the value obtained from Table 3.5.

3.27 Calculate the Clebsch-Gordan coefficient $C^{j_a\ 1\ j_a+1}_{j_a\ 1\ j_a+1}$ using (3.5.59) and verify that the answer agrees with the value obtained from Table 3.5.

3.28 Calculate the Clebsch-Gordan coefficient $C^{\ell\quad \frac{1}{2}\ \ \ell-\frac{1}{2}}_{\ell-1\ \frac{1}{2}\ \ \ell-\frac{1}{2}}$ using (3.5.60) and and verify that the answer agrees with the value obtained from Table 3.4.

3.29 Calculate the Clebsch-Gordan coefficient using $C^{j_a\quad 1\ \ m}_{m-1\ 1\ \ m}$ using (3.5.60) and verify that the answer agrees with the value obtained from Table 3.5 for each of the three cases $m = j_a + 1$, $m = j_a$ and $m = j_a - 1$.

3.30 Calculate the Clebsch-Gordan coefficient using $C^{j_a\quad\ 1\ \ j_a+1}_{m+1\ -1\ \ m}$ using (3.5.62) and verify that the answer agrees with the value obtained from Table 3.5. Which values of the index n must be included in the sum?

3.31 Calculate the Clebsch-Gordan coefficient using $C^{j_a\quad\ 1\ \ j_a}_{j_a-1\ 0\ \ j_a-1}$ using (3.5.62) and verify that the answer agrees with the value obtained from Table 3.5. Which values of the index n must be included in the sum?

For Sect. 3.6

3.32

(a) Show that (3.6.6) with $\kappa = 0$ agrees with (3.6.5a).
(b) Show that (3.6.6) with $\kappa = \pm 1$ agrees with (3.6.5b).

3.33

(a) Calculate the reduced matrix element $\langle j' \| T^{(J)} \| j \rangle$ for $T^{(J)}_{\kappa=+1} = J_{+1}$.
(b) Calculate the reduced matrix element $\langle j' \| T^{(J)} \| j \rangle$ for $T^{(J)}_{\kappa=-1} = J_{-1}$.

3.34 Using the explicit expressions for c_j, a_j, and d_j given in (3.6.19), show that (3.6.20) is identical to (3.6.18).

3.35

(a) Taking $\kappa = +1$ in (3.6.17) and using the explicit expressions for c_j, a_j, and d_j given in (3.6.19), derive (3.6.21).
(b) Taking $\kappa = -1$ in (3.6.17) and using the explicit expressions for c_j, a_j, and d_j given in (3.6.19), derive (3.6.22).

3.36

(a) When V_κ is hermitian show that in (3.6.20)–(3.6.22) the functions a_j and c_j satisfy $(a_j)^* = a_j$ and $(c_j)^* = -c_j$. Be certain to express d_{j-1} in terms of c_j.
(b) When V_κ is skew-hermitian show that in (3.6.20)–(3.6.22) the functions a_j and c_j satisfy $(a_j)^* = -a_j$ and $(c_j)^* = c_j$.

Chapter 4
Stationary Perturbation Theory

4.1 Stationary Perturbation Theory and Its Underlying Assumptions

4.1.1 Introduction

The description of any physical system is only approximate. It is true, of course, that it is possible to solve some models exactly, but such models in turn only approximately describe real physical systems. For example, the harmonic oscillator describes a vibrating diatomic molecule only when the vibrations are not too violent and anharmonic forces are negligible. In a similar vein, it is possible to find the energy levels of the hydrogen atom using the appropriate non-relativistic Hamiltonian. But the non-relativistic Hamiltonian or Schrödinger equation is only approximate: to take into account the spin of the electron, the Pauli equation or Pauli Hamiltonian must be used to describe the electron. The Dirac equation is a relativistic equation that describes a spin-1/2 particle. To take into account the fact that electromagnetic signals travel between the proton and electron at the speed of light instead of instantaneously, the Bethe-Salpeter equation is employed. Even when the electron is traveling at speeds far less than that of light, the electron interacts electromagnetically with itself, causing small but experimentally detectable deviations from the energy levels calculated from the Pauli equation or Pauli Hamiltonian. Problems in physics cannot usually be solved exactly. Therefore, if a physicist wishes to calculate numbers that can be compared with experimental values, approximations are inevitable.

The most frequently used approximation method is based on the splitting of the Hamiltonian H into two parts,

$$H = H_0 + H_1 \,. \tag{4.1.1}$$

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9_4

This splitting is done so that H_0, often called the free or interaction-free Hamiltonian, is the Hamiltonian of a system such as an oscillator or a rotator with an eigenvalue equation that can be solved exactly. The Hamiltonian H_1 describes some small interaction that slightly "perturbs" the interaction-free system.

The physical system with the "exact" Hamiltonian H is assumed to be similar to a model system with a Hamiltonian H_0. Here the word "similar" means that each energy eigenvalue and each energy eigenvector of H corresponds to one energy eigenvalue and one energy eigenvector of H_0. This correspondence is such that when $H_1 \rightarrow 0$, the energy eigenvalues and energy eigenvectors of H go into those of the model system described by H_0. While an exactly solvable, similar system does not exist for every physical system, there are numerous cases for which similar systems do exist: hydrogen-like atoms and systems in an external field are examples. Perturbation theory offers one of the most powerful tools for studying such systems.

4.1.2 Magnetic Moment in a Magnetic Field

The interaction-free Hamiltonian H_0 is chosen in this case to be the rotator Hamiltonian, $H_0 = \mathbf{J}^2/2I$ of (2.2.30). To conjecture the interaction Hamiltonian H_1 for the interaction with an external magnetic field $\mathbf{B}$, the usual procedure is followed by starting with the energy of the corresponding classical system and making the transition to the quantum system by replacing the classical observables by their corresponding quantum operators. The magnetic field acts on a magnetic dipole moment $\boldsymbol{\mu}$. The potential energy U of a classical dipole moment $\boldsymbol{\mu}$ in a magnetic field $\mathbf{B}$ is given by

$$U = -\boldsymbol{\mu} \cdot \mathbf{B}\,. \tag{4.1.2}$$

Example 4.1.1 Show that the magnetic dipole moment $\boldsymbol{\mu}$ of a mass-point with mass m and electric charge q moving in a circular orbit with orbital angular momentum $\mathbf{l}$ is given by

$$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{l}\,. \tag{4.1.3}$$

Solution Denoting the radius of the orbit by R and the period by T, the speed v of the mass-point is given by $v = 2\pi R/T$. Recalling that the magnetic dipole moment is the product of the current $I = q/T$ and the area πR^2 of the orbit,

$$\mu = \frac{q}{T}\pi R^2 = \frac{qR}{2}\frac{2\pi R}{T} = \frac{qRv}{2} = \frac{q}{2m}Rmv = \frac{q}{2m}l\,,$$

where the angular momentum $l = Rmv$. Since $\mathbf{l}$ and $\boldsymbol{\mu}$ are in the same direction when q is positive and in opposite directions when q is negative, (4.1.3) immediately follows.

In (4.1.3) two masses m_e and $m_p \approx m_n \approx 1836m_e$ commonly occur that are of different orders of magnitude. Consequently in quantum mechanics two units for a magnetic dipole moment are often used:

(1) The Bohr magneton μ_B is obtained from (4.1.3) by making the replacements $q \to e, m \to m_{\text{electron}} \equiv m_e$ and $\mathbf{l} \to \hbar$:

$$\text{Bohr magneton} \quad \mu_B \equiv \frac{e\hbar}{2m_e} = 9.27 \times 10^{-24} \text{ J/ Tesla} \tag{4.1.4}$$

(2) The nuclear Bohr magneton μ_N, the second commonly used unit, is identical to the Bohr magneton except that the electron's mass is replaced by the proton's mass:

$$\text{Nuclear Bohr magneton} \quad \mu_N \equiv \frac{e\hbar}{2m_p} = 5.049 \times 10^{-27} \text{ J/ Tesla} \tag{4.1.5}$$

The magnetic moment of a rotating or spinning physical object with angular momentum $\mathbf{j}$ is, in general, given by

$$\boldsymbol{\mu} = g\frac{q}{2m}\mathbf{j}\,. \tag{4.1.6}$$

Here q is the charge of the object and the factor g, called the Landé factor or gyromagnetic ratio, is a parameter that takes a specific value characteristic of the particular physical object. If the physical system is a classical relativistic spinning particle with no intrinsic structure and with angular momentum $\mathbf{s}$ about its center of mass, then it is possible to show that $g = 2$. The electron with $q = -e$ is a nearly structureless, relativistic particle, and its Landé factor satisfies $g_{\text{electron}} \cong 2$.[1] Thus

$$\boldsymbol{\mu}_{\text{electron}} = g_{\text{electron}} \left(\frac{-e}{2m_e}\right) \mathbf{s} \cong 2 \left(\frac{-e}{2m_e}\right) \mathbf{s}\,, \tag{4.1.7}$$

where $\mathbf{s}$ is the spin of the electron.

The magnetic moment of a proton, an extended, relativistic object with intrinsic structure, has been measured to be

$$\boldsymbol{\mu}_{\text{proton}} = (2 + 3.586)\frac{e}{2m_p}\mathbf{s}\,. \tag{4.1.8a}$$

The magnitude of the proton's spin is $\hbar/2$ so

$$\mu_{\text{proton}} = (2 + 3.586)\frac{e}{2m_p}\frac{\hbar}{2} = 2.793\,\mu_N\,. \tag{4.1.8b}$$

---

[1] More precisely, as a result of corrections calculated from quantum electrodynamics, $g_{\text{electron}} = 2[1 + \frac{\alpha}{2\pi} - 0.328\left(\frac{\alpha}{2\pi}\right)^2 + \cdots]$ where $\alpha = e^2/(4\pi\epsilon_0\hbar c) \cong 1/137.036$ is the Sommerfeld constant.

The Landé factor for the proton is $g_{\text{proton}} = 2 + 3.586$ where $3.586[e/(2m_p)]\,\mathbf{s}$ is called the *anomalous* magnetic moment. It is due to the intrinsic structure of the proton, which, unlike the electron, is not a nearly structureless particle.

A neutron has structure but no net charge. Its magnetic moment is, therefore, only anomalous and is empirically determined to be

$$\boldsymbol{\mu}_{\text{neutron}} = (0 - 3.826)\frac{e}{2m_n}\mathbf{s}\,, \tag{4.1.9a}$$

implying that

$$\mu_{\text{neutron}} = (0 - 3.826)\frac{e}{2m_n}\frac{\hbar}{2} = -1.913\,\mu_N\,. \tag{4.1.9b}$$

Because the Bohr magneton is three orders of magnitude greater than the nuclear Bohr magneton, the small nuclear magnetic moments of a proton or a nucleus give noticeable effects only when the much larger magnetic moments of electrons are absent.

In quantum mechanics the magnetic dipole moment operator μ is obtained by replacing the classical angular momenta $\mathbf{j}$ by the angular momentum operators $\mathbf{J}$:

$$\boldsymbol{\mu} = g\frac{q}{2m}\mathbf{J}\,, \tag{4.1.10}$$

where the J_i fulfill the algebra of angular momentum (2.3.1). The J_i can be the orbital angular momentum L_i, the spin angular momentum S_i, or the sum of spin and orbital angular momentum. For example, for an orbiting, spinning electron, the magnetic dipole moment operator is

$$\boldsymbol{\mu} = \boldsymbol{\mu}_{\text{orbital}} + \boldsymbol{\mu}_{\text{spin}} = -\frac{e}{2m_e}(\mathbf{L} + 2\mathbf{S})\,. \tag{4.1.11}$$

When the physical system is a rotating molecule, a magnetic dipole moment is produced by the rotational motion of the entire molecule: these magnetic moments result from the rotation of the positively charged nuclei about the center of mass and the rotation of the negatively charged electron clouds that move with the nuclei. The rotation of the charge distribution results in a magnetic moment that, as before, is proportional to the angular momentum of the molecule and inversely proportional to the mass of the nucleus,

$$\boldsymbol{\mu}_{\text{molecule}} = g\frac{q}{2M_{\text{nucleus}}}\mathbf{J}\,. \tag{4.1.12}$$

The Landé factor g is a measure of the current distribution of the rotating molecule and is determined empirically from the Zeeman splitting of the energy levels. (For example, for the diatomic molecule H_2, g = 0.883.)

The magnetic dipole moment produced by the rotational motion of the whole molecule is of the same order of magnitude as that of the magnetic moments of the nuclei. In their ground state, most molecules have no angular momentum that results from motion of the electrons. For these molecules the magnetic dipole moment is produced by the nuclei of the molecule or by the rotational motion of the entire molecule and is given by

$$\boldsymbol{\mu} = g\frac{e}{2m_p}\mathbf{J} = g\mu_N\frac{1}{\hbar}\mathbf{J}\,, \tag{4.1.13}$$

which is on the order of the magnitude of the nuclear Bohr magneton (4.1.5).

4.1.3 *Quantum Rotator in a Uniform Magnetic Field*

The quantum rotator in a classical magnetic field is a very special case in that it can be solved without perturbation theory, but it is nevertheless instructive to first study this case. If a quantum system with a magnetic moment $\boldsymbol{\mu}$ and a Hamiltonian $H_0 = \mathbf{J}^2/2I$ as given in (2.2.28) for a rotator is placed in a uniform magnetic field $\mathbf{B}$, from correspondence with the classical energy (4.1.2), the Hamiltonian for the quantum rotator in a magnetic field is

$$H = H_0 + H_1 = \frac{\mathbf{J}^2}{2I} - \boldsymbol{\mu}\cdot\mathbf{B} = \frac{\mathbf{J}^2}{2I} - g\frac{q}{2m}\mathbf{J}\cdot\mathbf{B}\,. \tag{4.1.14}$$

This Hamiltonian is no longer rotationally invariant, which means that it does not commute with angular momentum:

$$[H, J_i] = [\boldsymbol{\mu}, J_i]\cdot\mathbf{B} = g\frac{q}{2m}[J_k, J_i]\mathbf{B}_k \neq 0\,. \tag{4.1.15}$$

The magnetic field $\mathbf{B}$ distinguishes a direction and destroys the isotropy of space: the interaction Hamiltonian H_1 due to the magnetic field $\mathbf{B}$ has broken the rotational symmetry.

According to (4.1.14) the free or “unperturbed” Hamiltonian is given by

$$H_0 = \frac{\mathbf{J}^2}{2I}\,, \tag{4.1.16}$$

and the perturbation or interaction Hamiltonian is

$$H_1 = -g\frac{q}{2m}\mathbf{B}\cdot\mathbf{J}\,. \tag{4.1.17}$$

Perturbation theory is not needed to determine the energy levels of the Hamiltonian (4.1.14) provided the 3-direction is chosen to point in the direction of the magnetic field $\mathbf{B}$. That is, when a coordinate system is chosen so that $\mathbf{B} = (0, 0, B)$, then (4.1.14) becomes

$$H = \frac{1}{2I} J^2 - g \frac{q}{2m} B J_3 \,. \tag{4.1.18}$$

The eigenvectors of the Hamiltonian $H = H_0 + H_1$ are the eigenvectors $|j, j_3\rangle$ of the complete set of commuting observables J^2 and J_3 discussed in Chap. 3, Sect. 3.3. The eigenvalues E_j^0 of the free Hamiltonian H_0 are

$$H_0 |j, j_3\rangle = \frac{\mathbf{J}^2}{2I} |j, j_3\rangle = \frac{\hbar^2 j(j+1)}{2I} |j, j_3\rangle = E_j^0 |j, j_3\rangle \,. \tag{4.1.19}$$

The eigenvalues E_{j,j_3} of the exact Hamiltonian H are also immediately obtained by applying H as given in (4.1.18) to $|j, j_3\rangle$,

$$H|j, j_3\rangle = \left[\frac{\mathbf{J}^2}{2I} - g\frac{q}{2m} B J_3\right] |j, j_3\rangle = \left[\frac{\hbar^2 j(j+1)}{2I} - g\frac{q}{2m} B\hbar j_3\right] |j, j_3\rangle = E_{j,j_3} |j, j_3\rangle \,. \tag{4.1.20}$$

The eigenvalues of H_0 depend only on j. Thus each eigenvalue E_j^0 of H_0 has a $(2j+1)$-fold degeneracy, implying that to each E_j^0 there correspond the $(2j+1)$-dimensional space $\Re^j$. In contrast, the eigenvalues of H depend on both j and j_3. To each eigenvalue E_{j,j_3} there corresponds one eigenvector $|j, j_3\rangle$ or, equivalently, the one-dimensional space $R^j_{j_3}$. Therefore, when the external magnetic field is switched off, $B \to 0$, the (2j+1)-non-degenerate energy levels E_{j,j_3} of H merge into the (2j+1)-degenerate energy levels E_j^0 of H_0:

$$\begin{pmatrix} E_{j,j_3=j} \\ E_{j,j_3=j-1} \\ \vdots \\ E_{j,j_3=-j} \end{pmatrix} \to E_j^0 \quad \text{as} \quad B \to 0 \,. \tag{4.1.21}$$

The splitting of the energy levels in a magnetic field is called Zeeman splitting (or the Zeeman effect), and it occurs for atoms, molecules and any other rotationally invariant system that is placed in a magnetic field that breaks the symmetry.

If the Hamiltonian (4.1.14) had been used instead of (4.1.18), with $\mathbf{B}$ pointing in any direction other than along the 3-axis, the eigenvectors of J_3 could not be the eigenvectors of H. Eigenvectors $|j, j_n\rangle$ of H as given in (4.1.14) are eigenvectors of the complete set of commuting observables $\mathbf{J}^2$ and $\hat{\mathbf{n}} \cdot \mathbf{J}$, where $\hat{\mathbf{n}}$ is the unit vector in the direction of the magnetic field. Eigenvalues of H as given in (4.1.14) could then be found by proceeding as before using $|j, j_n\rangle$ instead of $|j, j_3\rangle$ where j_n is the eigenvalue of $\hat{\mathbf{n}} \cdot \mathbf{J}$.

The free and exact Hamiltonians H_0 and H given in (4.1.16) and (4.1.14), respectively, do not represent the most general case for free and exact Hamiltonians in perturbation theory because they commute, $[H, H_0] = 0$, a condition that is not in general fulfilled. In the general case the exact and free Hamiltonian do not commute, $[H, H_0] \neq 0$, as is the case for atoms and molecules in a magnetic field, for which H_0 would not be a function $f(\mathbf{J}^2)$. But H_0 could be a spherically symmetric, which implies that it satisfies the condition $[H_0, J_i] = 0$. Such systems will also have Zeeman splitting: although the physical situation is very similar to the one just considered, the eigenvalues E_{j,j_3} cannot be calculated exactly as in (4.1.20). If H and H_0 do not commute, the procedure called perturbation theory must be used. This general case is now considered.

4.1.4 Perturbation Theory

Approximate solutions for an "exact" Hamiltonian $H = H_0 + H_1$ can be obtained using a method of calculation called perturbation theory provided there exists a "free" Hamiltonian H_0 for which energy eigenvectors are both known and are also the best approximations to the energy eigenvectors of H. At first it might seem that if all the eigenvectors of H_0 are known, then the eigenvector of H_0 that is the best approximation to each energy eigenvector of H would automatically be known. Unfortunately this is not the case. If the energy eigenvalues of H_0 are degenerate, as they usually are, the linear combinations of the eigenvectors of H_0, all with the same energy E^0, that are the best approximations to eigenvectors of H with energy eigenvalues almost equal to E^0 are not necessarily known.

In addition to the energy quantum number, other quantum numbers are typically required to completely specify the state except for very special cases such as the one-dimensional harmonic oscillator. For example, the quantum number j_3 is, in addition to j, required to specify the states of a rotator.

The set of operators that, together with H_0, forms a complete set of commuting observables with known eigenvalues is denoted by $B_1, B_2, \cdots, B_{N-1}$, or collectively by B. Their respective eigenvalues $b_1, b_2, \cdots, b_{N-1}$ are collectively denoted by b, and the $b_1, b_2, \cdots, b_{N-1}$ are assumed to be known. As an example, for the Hydrogen Hamiltonian, $b_1 = \ell$ and $b_2 = \ell_3 = m$.

Case I The characteristic of the first case to be considered is that the same set of operators B also forms a complete set of commuting observables with H, $[H, B] = 0$. Then there exist two complete sets of commuting observables consisting of N operators,

$$\{H_0, B_1, B_2, \cdots, B_{N-1}\} \equiv \{H_0, B\}, \tag{4.1.22a}$$

$$\{H, B_1, B_2, \cdots, B_{N-1}\} \equiv \{H, B\} = \{H_0 + H_1, B\}. \tag{4.1.22b}$$

The solutions of the physical system described by (4.1.22a) are the known eigenvectors $|E_n^0, b\rangle \equiv |E_n^0, b_1, b_2, \cdots, b_{N-1}\rangle$ and known eigenvalues $E_{n,b}^0 \equiv E_{n,b_1,b_2,\cdots,b_{N-1}}^0$. The solutions of the physical system described by (4.1.22b), which are not known, are the eigenvectors $|E_{n,b}, b\rangle \equiv |E_{n,b}, b_1, b_2, \cdots, b_{N-1}\rangle$ and energy eigenvalues $E_{n,b} \equiv E_{n,b_1,b_2,\cdots,b_{N-1}}$. The eigenvalues of H_0 often are degenerate—they do not depend on b—while the eigenvalues of H usually are not degenerate—they usually depend on b. For example, for the rotator the eigenvalues of $H_0 = \mathbf{J}^2/2I$ depend only on j while the eigenvalues of H depend on both j and j_3.

In perturbation theory when $H_1 \to 0$, the situation is very similar to that depicted in (4.1.21). The energy levels E_n^0 of the interaction-free system are degenerate. To each of the energy levels E_n^0 there corresponds a finite or infinite number p of energy levels $E_{n,b}$ of the system with interaction H_1. When $H_1 \to 0$, all of the p energy levels $E_{n,b}$ merge into the same energy value E_n^0,

$$\begin{pmatrix} E_{n,b_1} \\ E_{n,b_2} \\ \vdots \\ E_{n,b_{N-1}} \end{pmatrix} \to E_n^0 \quad \text{as} \quad H_1 \to 0\,. \tag{4.1.23}$$

As a consequence, p eigenvalues of H correspond to one eigenvalue of H_0. The eigenvectors of H and H_0 satisfy

$$\lim_{H_1\to 0} |E_{n,b}, b\rangle = |E_n^0, b\rangle \quad \text{where} \quad b = b_1, b_2, \cdot\cdot\cdot, b_{N-1}\,. \tag{4.1.24}$$

As mentioned earlier, the key to obtaining solutions with perturbation theory is to determine the energy eigenvector of H_0 that is the best approximation to each energy eigenvector of H. Even when E_n^0 is degenerate, for the case that H_0 and H both commute with B, (4.1.24) reveals that the best approximation to the energy eigenvector $|E_{n,b}, b\rangle$ of H and B is the eigenvector $|E^0, b\rangle$ of H_0 and B that has the *same* quantum numbers b! For example, for the case of a spin-1/2 rotator in an external magnetic field, there are two eigenvectors of H_0 with the same energy eigenvalue $E_j^0 = \hbar^2(1/2)(1/2+1)/2I$. Namely, $|E_j^0, +1/2\rangle$ and $|E_j^0, -1/2\rangle$. If, for example, an eigenvector of H is desired with $j = 1/2$, $j_3 = 1/2$, the eigenvector $|E_j^0, 1/2\rangle$ of H_0 with $j_3 = 1/2$ is the best approximation because it is the only eigenvector of H_0 with the desired eigenvalue $j_3 = 1/2$.

In the typical situation for which perturbation theory applies, there are p *known* eigenvectors $|E_n^0, b\rangle$—where p can be either a finite number or infinite—belonging to the one *known* eigenvalue E_n^0 of H_0. The p eigenvectors span the finite or infinite-dimensional space $\mathfrak{R}_0^p$ spanned by the eigenvectors of H_0 with the same eigenvalue E_n^0. The perturbation H_1 then splits the energy level E_n^0 into the sublevels $E_{n,b}$, removing the degeneracy of the eigenvalue E_n^0. Often—as is the case for the rotator in a magnetic field—the degeneracy stems from symmetries in the unperturbed Hamiltonian H_0. The perturbation Hamiltonian H_1 then breaks the symmetry, thereby removing the degeneracy.[2] The perturbation Hamiltonian H_1 also transforms the space $\mathfrak{R}^p_0$, in the whole space $\mathscr{H} = \sum_n \oplus \mathfrak{R}^n_0$, into another space $\mathfrak{R}^p$ spanned by the $|E_{n,b}, b\rangle$.

In order for (4.1.23) and (4.1.24) to make physical sense, the splitting inside a multiplet (between the $E_{n,b}$ for a given value n and different values of b) should be small compared to the splitting between different multiplets (between the $E_{n,b}$ for a given value b and different values of n). Thus the perturbation Hamiltonian H_1 must be "small" in a certain mathematical sense (e.g. $|\langle E^0_n, b|H_1|E^0_{n'}, b'\rangle| \ll |E^0_{n'} - E^0_n|$). How to precisely formulate the mathematical conditions on the operators H_0 and H such that (4.1.23) and (4.1.24) result and such that the perturbation theory converges is a difficult mathematical problem. Physicists use perturbation calculations to evaluate numbers and hope that H_1 has been chosen in such a way that the perturbation theory somehow converges.

When a situation arises such that $[H_0, B] = 0$ but $[H, B] \neq 0$, the label b could be used to label the eigenvectors of H, even though the eigenvectors of H are not eigenvectors of the operators B, revealing that the label b in $|E_{n,b}, b\rangle$ is only an "approximate" quantum number defined by the limit of (4.1.24). Such a procedure is often followed for the continuous eigenvalues E of H in scattering theory where the b are the asymptotic momenta $\mathbf{p}$ and $E^0 =$ eigenvalue of $H_0 = \mathbf{p}^2/2m$.

When H and H_0 have discrete spectra, rather than using eigenvectors labeled by b, it is preferable to find a system of operators $A_1, A_2, \cdots, A_{N-1}$ that, together with H, forms a complete set of commuting observables

$$[H, A] = 0\,, \tag{4.1.25}$$

and also has the property that

$$[H_0, A] = 0\,. \tag{4.1.26}$$

Then a new system of basis eigenvectors $|E^0_n, a\rangle$ of the interaction-free system, which are eigenvectors of the complete set of commuting observables $\{H_0, A\}$ are obtained by a linear transformation from the known interaction-free basis eigenvectors $|E^0_n, b\rangle$:

$$|E^0_n, a\rangle = \sum_b |E^0_n, b\rangle\langle E^0_n, b|E^0_n, a\rangle\,. \tag{4.1.27}$$

---

[2] The one-dimensional harmonic oscillator is a special case of little practical interest for which there is no degeneracy so that there is only one eigenvector $|E^0_n\rangle$ of H_0 for each value of E^0_n. Then no splitting can occur and there is only one eigenvalue E_n of H that is shifted slightly from E^0_n. This case is usually discussed in most textbooks under the heading of "non-degenerate perturbation theory".

The new basis vectors $|E_n^0, a\rangle$ are eigenvectors of H_0 with eigenvalue E_n^0 as are the $|E_n^0, b\rangle$. The exact eigenvectors $|E, a\rangle$, which are eigenvectors of the complete set of commuting observables $\{H, A\} = \{H, A_1, A_2, \ldots, A_{N-1}\}$ are then obtained from the known $|E_n^0, a\rangle$ using perturbation theory. Thus for the case that $[H, B] \neq 0$, but $[H_0, B] = 0$, perturbation for a degenerate energy spectrum has been reduced by the linear transformation in (4.1.27) to the previous case (4.1.22) that can be treated as if there were no degeneracy. The transformation matrix, $\langle E_n^0, b|E_n^0, a\rangle$ depends on the properties of the set of operators $B = B_1, B_2, \ldots, B_{N-1}$ and $A = A_1, A_2, \ldots, A_{N-1}$ and not on H and H_0.

An example of the above situation is the complete set of commuting observables H_0, B given by

$$H_0, \mathbf{L}^2, L_3, \mathbf{S}^2, S_3\,, \tag{4.1.28}$$

where $\mathbf{L}$ is the orbital angular momentum, $\mathbf{S}$ is the spin, and H_0 is a function of the orbital observables only and not of the spin. The eigenvectors $|E_n^0, b)$ of the set of operators (4.1.28) are the direct-product basis vectors

$$|E_n^0, b\rangle = |E_n^0, \ell, \ell_3, s, s_3, \rangle = |E_n^0, \ell, \ell_3\rangle \otimes |s, s_3\rangle\,, \tag{4.1.29}$$

where $s = 1/2$ and $\ell, = 0, 1, 2, \cdots$.

The exact Hamiltonian H contains an $\mathbf{L} - \mathbf{S}$ coupling and is given by,

$$H = H_0 + f\mathbf{L}\cdot\mathbf{S}\,, \tag{4.1.30}$$

where f may or may not be a function of operators that commute with L_i and S_i. The function f may, for example, be a function of $\mathbf{Q}^2$, $f = f(\mathbf{Q}^2)$, where the Q_i are the components of the position operator $\mathbf{Q}$. The exact Hamiltonian H will not commute with L_i and S_i because

$$[H, L_i] \neq 0\,, \quad [H, S_i] \neq 0\,. \tag{4.1.31}$$

However, the exact Hamiltonian will commute with the total angular momentum operator $\mathbf{J} = \mathbf{L} + \mathbf{S}$,

$$[H, J_i] = 0\,. \tag{4.1.32}$$

The above result is expected because the Hamiltonian is rotationally invariant, but it also follows immediately from the identity

$$\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2)\,, \tag{4.1.33}$$

and the fact that H_0 commutes with J_i. Therefore,

$$\{H_0, A\} = \{H_0, \mathbf{J}^2, J_3, \mathbf{L}^2, \mathbf{S}^2\}\,, \quad \text{where} \quad A = \mathbf{J}^2, J_3, \mathbf{L}^2, \mathbf{S}^2 \quad \text{and} \tag{4.1.34a}$$

$$\{H, A\} = \{H_0, \mathbf{J}^2, J_3, \mathbf{L}^2, \mathbf{S}^2\} \tag{4.1.34b}$$

are complete sets of commuting observables that differ only in the operators H and H_0.

The eigenvectors of the complete set of commuting observables (4.1.34b),

$$|E_n^0, a\rangle = |E_n^0, j, j_3, \ell, s\rangle\,, \tag{4.1.35}$$

are obtained by expanding them in terms of direct-product vectors (4.1.29),

$$|E_n^0, j, j_3, \ell, s\rangle = \sum_{\ell_3, s_3} |E_n^0, \ell, \ell_3, s, s_3\rangle\langle E_n^0, \ell, \ell_3, s, s_3|E_n^0, j, j_3, \ell, s\rangle\,. \tag{4.1.36a}$$

In the above equation E_n^0 depends only on the variable n, allowing the two vectors in the scalar product to be written as

$$|E_n^0, j, j_3, \ell, s\rangle = \sum_{\ell_3, s_3} |E_n^0, \ell, \ell_3, s, s_3\rangle[\langle E_n^0| \otimes \langle \ell, \ell_3, s, s_3|][|E_n^0\rangle \otimes |j, j_3, \ell, s\rangle]\,. \tag{4.1.36b}$$

For normalized energy eigenstates, $\langle E_n^0|E_n^0\rangle = 1$, yielding the desired result,

$$|E_n^0, j, j_3, \ell, s\rangle = \sum_{\ell_3, s_3} |E_n^0, \ell, \ell_3, s, s_3\rangle\langle \ell, \ell_3, s, s_3|j, j_3, \ell, s\rangle\,, \tag{4.1.36c}$$

where $\langle \ell, \ell_3, s, s_3|j, j_3, \ell, s\rangle \equiv \langle \ell, \ell_3, s, s_3|j, j_3\rangle$ are the Clebsch Gordon coefficients discussed in Chap. 3, Sect. 3.5.

Case II There is a second very special case in which it is possible to immediately determine the energy eigenvector of H_0 that is the best approximation to a specific eigenvector of H. If the energy eigenvalues of H_0 are not degenerate, so that there is only one eigenvector $|E_n^0, b_i\rangle$ of H_0 for each value of E_n^0, then the *single* energy eigenvector of H_0 with energy E_n^0 is the best approximation to an eigenvector of H with almost the same energy. For example, for a rotator in constant magnetic fields B_3 and B_1 that point in the 3- and 1-directions, respectively, the Hamiltonian H is given by

$$H = \frac{\mathbf{J}^2}{2I} - g\frac{q}{2m}B_3 J_3 - g\frac{q}{2m}B_1 J_1\,. \tag{4.1.37}$$

If $B_1 \ll B_3$, the last term can be treated as a perturbation:

$$H_0 = \frac{\mathbf{J}^2}{2I} - g\frac{q}{2m}B_3 J_3\,, \qquad H_1 = -g\frac{q}{2m}B_1 J_1\,. \tag{4.1.38}$$

The eigenvector of H_0 has an energy equal to

$$E^0_{j,j_3} = \frac{\hbar^2 j(j+1)}{2I} - g\frac{q}{2m}B_3 j_3 \tag{4.1.39}$$

and its eigenvector is the *single* eigenvector $|E^0_{j,j_3}, j_3\rangle = |j, j_3\rangle$. So even though there exists no operator B that commutes with both H_0 and H, it is possible to immediately identify the eigenvector of H_0 that is the best approximation to any eigenvector of H.

The very special case in which there is no degeneracy in the eigenvalues of H_0, and there is only one eigenvalue E_n of H that is slightly shifted from E^0_n will be considered as will the case where there exists an operator B that commutes both with H_0 and H. The two cases need not be considered separately because in either case the energy eigenvector of H_0 is already known that is the best approximation to each eigenvector of H. Specifically, in the next section perturbation theory will be developed that determines the unknown eigenvalues E_n, b_i and eigenvectors $|E_{n,b_i}, b_i\rangle$ of H in terms of the known eigenvectors $|E^0_n, b_i\rangle$ of H_0.

Case III One final case remains that requires special consideration. If an operator B cannot be found that commutes with both H_0 and H, and if the energy eigenvalues of H_0 are degenerate, then the eigenvector of H_0 that is the best approximation to a specific eigenvector of H is not immediately known. For example, assume that H_0 and B form a complete set of commuting operators, and H and C $(B \neq C)$ also form a complete set of commuting operators. If E^0_n is doubly degenerate, there are two eigenvectors of H_0 that have the eigenvalue E^0_n, namely $|E^0_n, b_1\rangle$ and $|E^0_n, b_2\rangle$. In general the eigenvector of H_0 that is the best approximation to an eigenvector $|E_{n,c_1}, c_1\rangle$ of H with an energy eigenvalue E_{n,c_1} that almost equals E^0_n is a linear combination of the two eigenvectors of H_0 with energy E^0_n,

$$|E_{n,d_1}\,, d_1\rangle \cong \alpha_1|E^0_n\,, b_1\rangle + \alpha_2|E^0_n\,, b_2\rangle \tag{4.1.40}$$

where the α_i are constants that must be determined. So for this final case, which will be discussed in Sect. 4.3, the eigenvector of H_0 that is the best approximation to each eigenvector of H is not immediately known but instead must be determined by solving for the α_i.

4.2 Stationary Perturbation Expansion

In this section approximate formulas are derived for the energy eigenvalues and energy eigenvectors of $H = H_0 + H_1$ in terms of the energy eigenvectors and energy eigenvalues of H_0. The formulas derived here can be used for the cases listed below:

Case I A complete set of hermitian operators B is known that satisfies

$$[B, H_0] = [B, H] = 0\,. \tag{4.2.1}$$

Case II The eigenvalues of H_0 are non-degenerate.

The second case is a special situation in which the eigenvalues b of B take only a single value. The third and final case will be considered in the next section:

Case III The energy eigenvalues of the free Hamiltonian H_0 are degenerate and no complete set of hermitian operators B is known that commutes with both H_0 and H.

In all three cases the free or unperturbed Hamiltonian H_0 is assumed to have been solved in a basis $|E_n^0, b\rangle$ satisfying

$$H_0|E_n^0, b\rangle = E_n^0|E_n^0, b\rangle\,, \tag{4.2.2a}$$

$$B|E_n^0, b\rangle = b|E_n^0, b\rangle\,. \tag{4.2.2b}$$

Usually the eigenvalues of H_0 depend only on the quantum number n and not on b although there are exceptions. For later convenience the eigenvectors $|E_n^0, b\rangle$ are normalized as follows:

$$\langle E_{n'}^0, b'|E_n^0, b\rangle = \delta_{n,n'}\, \delta_{b,b'}\,. \tag{4.2.3}$$

For Case I eigenvectors $|E_{n,b}, b\rangle$ of the "exact" Hamiltonian H are sought that satisfy

$$H|E_{n,b}, b\rangle = E_{n,b}|E_{n,b}, b\rangle\,, \tag{4.2.4a}$$

$$B|E_{n,b}, b\rangle = b|E_{n,b}, b\rangle\,, \tag{4.2.4b}$$

where

$$|E_{n,b}, b\rangle \underset{H_1 \to 0}{\longrightarrow} |E_n^0, b\rangle\,. \tag{4.2.5}$$

Even if the energy eigenvalues E_n^0 are degenerate, the eigenvector of H_0 that is the best approximation to $|E_{n,b}, b\rangle$ is known: it is the eigenvector of H_0 that has the same quantum numbers n and b.

For Case II eigenvectors $|E_n\rangle$ of the "exact" Hamiltonian H are sought that satisfy

$$H|E_n\rangle = E_n|E_n\rangle\,, \tag{4.2.6}$$

where

$$|E_n\rangle \underset{H_1\to 0}{\longrightarrow} |E_n^0\rangle\,. \tag{4.2.7}$$

If H commuted with some other operator, the eigenvectors $|E_n\rangle$ would be labeled by additional quantum numbers that are omitted here because they are of no consequence for the discussion. Because the energy eigenvalues of H_0 are nondegenerate, the eigenvector of H_0 that is the best approximation to $|E_n\rangle$ is the eigenvector of H_0 with the same quantum number n. For both Case I and Case II, approximate energy eigenvalues and eigenvectors are known. Case III must be considered separately because approximate energy eigenvectors are not initially known and must be calculated.

Cases I and II Having found the eigenvectors $|E_n^0, b\rangle$, the eigenvalue b plays a passive role in all remaining formulas. That is, the eigenvalue b is the same for all eigenvalues and eigenvectors in the formulas that follow in this section. Therefore, to simplify notation, the index b is suppressed by adopting the notation

$$E_{n,b} \to E_n\,, \quad E_n^0 \to E_n^0\,, \quad |E_{n,b}\,, b\rangle \to |E_n\rangle\,, \quad |E_n^0\, b\rangle \to |E_n^0\rangle\,, \tag{4.2.8}$$

while keeping in mind that the energy eigenvalues and eigenvectors do, in general, depend on b. (The reader can, of course, put the index b back into the formulas by reversing the above substitutions.) However, having suppressed the index b, the notation for the eigenvectors of H_0 is the same for Case I and Case II, so both cases can be considered simultaneously.

The perturbation calculation is begun by considering the following matrix element of (4.1.1):

$$\langle E_{n'}^0|H|E_n\rangle = \langle E_{n'}^0|H_0 + H_1|E_n\rangle\,. \tag{4.2.9}$$

The matrix element on the left-hand side of (4.2.9) is simplified using the fact that $|E_n\rangle$ is, by definition, an eigenvector of H with (unknown) eigenvalue E_n. The first matrix element on the right-hand side of (4.2.7) is evaluated using the Hermiticity of H_0 and (4.2.2a). Thus

$$E_n\langle E_{n'}^0|E_n\rangle = E_{n'}^0\langle E_{n'}^0|E_n\rangle + \langle E_{n'}^0|H_1|E_n\rangle\,. \tag{4.2.10}$$

Since the perturbation is assumed to be small, the energy eigenvalues E_n and E_n^0 do not differ by much. For the same reason the eigenvectors $|E_n\rangle$ and $|E_n^0\rangle$ are almost the same, implying that the matrix element $\langle E^0_{n'=n}|E_n\rangle \approx \langle E^0_n|E^0_n\rangle = 1$. Taking $n' = n$ in (4.2.10), it is thus possible to divide by the non-zero matrix element $\langle E^0_n|E_n\rangle$,

$$E_n = E^0_n + \frac{\langle E^0_n|H_1|E_n\rangle}{\langle E^0_n|E_n\rangle} = E^0_n + \Delta E_n \,. \tag{4.2.11}$$

Equation (4.2.11), which is one of two fundamental equations in stationary perturbation theory, reveals that the exact energy values E_n are equal to the energy values E^0_n of the interaction-free system plus a small perturbation ΔE_n. As can be seen from the above equation, ΔE_n is given by the matrix element of the "small" interaction Hamiltonian H_1 and must be calculated.

To derive the second fundamental equation of stationary perturbation theory, the "exact" eigenvectors $|E_n\rangle$ of H are expanded in terms of the known eigenvectors $|E^0_{n'}\rangle$ of H_0. The eigenvectors $|E^0_n\rangle$ of the complete set of commuting observables H_0 (and B) form a complete basis system in the space $\mathscr{H}$ of state vectors. Thus every vector $\phi \in \mathscr{H}$ can be expanded in terms of this basis system:

$$\phi = \sum_{n'} |E^0_{n'}\rangle\langle E^0_{n'}|\phi\rangle \,. \tag{4.2.12}$$

In particular each eigenvector $\phi = |E_n\rangle$ of the complete set of commuting observables H (and B) can be expanded as in (4.2.12)

$$|E_n\rangle = \sum_{n'} |E^0_{n'}\rangle\langle E^0_{n'}|E_n\rangle = |E^0_n\rangle\langle E^0_n|E_n\rangle + \sum_{n'\neq n} |E^0_{n'}\rangle\langle E^0_{n'}|E_n\rangle \,. \tag{4.2.13}$$

For a given value of n, the terms with $n' = n$ and $n' \neq n$ are of different magnitudes:

$$\langle E^0_n|E_n\rangle \approx \langle E^0_n|E^0_n\rangle = 1 \,, \tag{4.2.14a}$$

$$\langle E^0_{n'}|E_n\rangle \approx \langle E^0_{n'}|E^0_n\rangle = 0 \quad \text{for} \quad n' \neq n \,. \tag{4.2.14b}$$

Therefore, the eigenvector $|E_n\rangle$ of H is written as the corresponding eigenvector $|E^0_n\rangle$ plus an (infinite) sum of terms that are much smaller in magnitude. In (4.2.13) the large term $|E^0_n\rangle$ is isolated by adding and then subtracting it,

$$|E_n\rangle = |E^0_n\rangle + [\langle E^0_n|E_n\rangle - 1]|E^0_n\rangle + \sum_{n'\neq n} |E^0_{n'}\rangle\langle E^0_{n'}|E_n\rangle \,. \tag{4.2.15}$$

As a consequence of (4.2.14), all terms on the right-hand side of (4.2.15) are small except the first.

A more convenient expression for the final matrix element in (4.2.15) can be obtained by taking $n' \neq n$ in (4.2.10). The energy difference $E_n - E^0_{n'}$ is different from zero because $E_n - E^0_{n'} \approx E^0_n - E^0_{n'} \neq 0$. Therefore, it is possible to divide (4.2.10) by $E_n - E^0_{n'}$,

$$\langle E^0_{n'}|E_n\rangle = \frac{\langle E^0_{n'}|H_1|E_n\rangle}{E_n - E^0_{n'}} \qquad n' \neq n\,. \tag{4.2.16}$$

Using the above result (4.2.15) becomes

$$|E_n\rangle = |E^0_n\rangle + [\langle E^0_n|E_n\rangle - 1]|E^0_n\rangle + \sum_{n'\neq n} |E^0_{n'}\rangle \frac{\langle E^0_{n'}|H_1|E_n\rangle}{E_n - E^0_{n'}}\,. \tag{4.2.17}$$

The summation in (4.2.17) should actually be over all possible values of the quantum numbers n and b that characterize the state $|E^0_n, b\rangle$. However, as shown in the following example, the matrix element of H_1 in the above equation is zero unless $b = b'$ so it is not necessary to sum over the eigenvalue b. If the eigenvalue b were put back into (4.2.17), each eigenvector would just have the same eigenvalue b.

Example 4.2.1 If there exists an hermitian operator B that commutes with both H_0 and $H = H_0 + H_1$, show that

$$\langle E^0_n, b_i|H_1|E_{n,b_j}, b_j\rangle = 0 \ \text{ if } \ b_i \neq b_j\,,$$

where

$$B|E^0_n, b_i\rangle = b_i|E^0_n, b_i\rangle\,, \qquad B|E_{n,b_j}, b_j\rangle = b_j|E_{n,b_j}, b_j\rangle\,.$$

Solution Because B commutes with both H_0 and H,

$$[B, H_1] = [B, H - H_0] = [B, H] - [B, H_0] = 0\,,$$

it also commutes with H_1. To establish the desired result, the first equation above implies

$$\begin{aligned} 0 = \langle E^0_n, b_i|[H_1B - BH_1]|E_{n,b_j}, b_j\rangle &= \langle E^0_n, b_i[H_1b_j - b_iH_1]|E_{n,b_j}, b_j\rangle \\ &= (b_j - b_i)\langle E^0_n, b_i|H_1|E_{n,b_j}, b_j\rangle\,. \end{aligned}$$

The above equation requires that

$$\langle E^0_n, b_i|H_1|E_{n,b_j}, b_j\rangle = 0 \ \text{ if } \ b_i \neq b_j\,.$$

Equations (4.2.11) and (4.2.17) are the two fundamental equations in stationary perturbation theory, and are, respectively, exact equations for the energy E_n and eigenvectors $|E_n\rangle$. But the unknown quantity $|E_n\rangle$ appears on the right-hand side of both equations, and E_n occurs on the right-hand side of (4.2.17). To solve the equations, the left-hand sides are successively substituted into the right-hand sides. In carrying out this iterative procedure, it is important to keep in mind that H_1 is a small perturbation. Thus on the right-hand side of (4.2.11) the second term is small compared with the first. Similarly, on the right-hand side of (4.2.17), both the second term and the sum over n' have small coefficients compared with the first term.

Zeroth-Order Approximation

The zeroth-order approximation for the eigenvalues E_n and eigenvectors $|E_n\rangle$, denoted by $E_n^{(0)}$ and $|E_n^{(0)})$, respectively, are obtained by setting $H_1 = 0$. Thus,

$$E_n \to E_n^{(0)} \to E^0\,, \quad |E_n\rangle \to |E_n^{(0)}) \to |E_n^0\rangle\,. \tag{4.2.18}$$

First-Order Approximation

Since the terms on the right-hand side of (4.2.11) and (4.2.17) that involve E_n or $|E_n\rangle$ are small, the first-order approximations to E_n and $|E_n\rangle$ are obtained by substituting their zeroth-order approximations on the right-hand sides of (4.2.11) and (4.2.17), respectively. Denoting the first-order approximation for energy by $E_n^{(1)}$ and using (4.2.18), (4.2.11) becomes

$$E_n^{(1)} = E_n^0 + \frac{\langle E_n^0|H_1|E_n^0\rangle}{\langle E_n^0|E_n^0\rangle}\,. \tag{4.2.19}$$

The orthogonality relation (4.2.3) for the eigenvectors $|E_n^0\rangle$ yields the desired result,

$$E_n^{(1)} = E_n^0 + \langle E_n^0|H_1|E_n^0\rangle = E_n^0 + \Delta E_n^{(1)}\,. \tag{4.2.20}$$

To first order the splitting $\Delta E_n^{(1)}$ of the degenerate energy levels E_n^0 of the free or unperturbed Hamiltonian equals the matrix element of the perturbation H_1 between the unperturbed, known eigenstates $|E_n^0\rangle$.

Denoting the first-order approximation for energy eigenvectors by $|E_n^{(1)}\rangle$ and using (4.2.18), (4.2.17) becomes

$$|E_n^{(1)}\rangle = |E_n^0\rangle + \Big[\langle E_n^0|E_n^0\rangle - 1\Big]|E_n^0\rangle + \sum_{n'\neq n} |E_{n'}^0\rangle \frac{\langle E_{n'}^0|H_1|E_n^0\rangle}{E_n^0 - E_{n'}^0}\,. \tag{4.2.21}$$

From (4.2.3) the matrix element $\langle E_n^0|E_n^0\rangle = 1$, so the above equation takes the final form

$$|E_n^{(1)}\rangle = |E_n^0\rangle + \sum_{n'\neq n} |E_{n'}^0\rangle \frac{\langle E_{n'}^0|H_1|E_n^0\rangle}{E_n^0 - E_{n'}^0}\,. \tag{4.2.22}$$

Example 4.2.2 Using first-order perturbation theory, calculate the ground-state energy level of the hydrogen atom when it experiences a perturbation $H_1 = g(Q_x^2 + Q_y^2 + Q_z^2)$ where g is a constant, and $Q_x = Q_{2x} - Q_{1x}$ etc. is the relative position as given in (3.2.31). (A force that increases linearly with the radial distance between the electron and proton yields a perturbation of the above form.)

Solution From Table 3.1, the normalized, ground-state wave function for hydrogen is

$$\langle \mathbf{x}|E_1^0\rangle \equiv \langle \mathbf{x}|E_1,\ \ell = 0,\ m = 0\rangle \equiv \psi_{1,0,0}(\mathbf{x}) = \frac{1}{(r_B)^{3/2}}\,\frac{1}{\sqrt{\pi}}\ e^{-\frac{r}{r_0}}\ ,$$

where $r_0 = 4\pi\epsilon_0\hbar^2/(\mu e^2)$. Because both the "free" and "exact" Hamiltonians H_0 and H commute with $\mathbf{L}^2$ and L_3, the eigenstates of both Hamiltonians can be labeled by the quantum numbers $b = \ell, m$. Since the additional quantum numbers $b = \ell, m$ have been suppressed, in the notation of this section, $\langle \mathbf{x}|E_1, \ell = 0, m = 0\rangle = \langle \mathbf{x}|E_1^0\rangle$. From (4.2.20), to first order in the perturbation, the ground-state energy is

$$E_1^{(1)} = E_1^0 + \langle E_1^0|H_1|E_1^0\rangle = E_1^0 + \langle E_1^0|g(Q_x^2 + Q_y^2 + Q_z^2)|E_1^0\rangle$$

Using (3.2.20) twice,

$$E_1^{(1)} = E_1^0 + \iint_{-\infty}^{\infty} \mathrm{d}^3x'\mathrm{d}^3x\,\langle E_1^0|\mathbf{x}'\rangle\langle \mathbf{x}'|g(Q_x^2 + Q_y^2 + Q_z^2)|\mathbf{x}\rangle\langle \mathbf{x}|E_1^0\rangle\,.$$

Recalling that the generalized position eigenvectors satisfy $Q_i|\mathbf{x}\rangle = x_i|\mathbf{x}\rangle$ and fulfill the normalization condition $\langle \mathbf{x}'|\mathbf{x}\rangle = \delta^3(\mathbf{x} - \mathbf{x}')$,

$$\begin{aligned}
E_1^{(1)} - E_1^0 &= \iint_{-\infty}^{\infty} \mathrm{d}^3x\,\mathrm{d}^3x'\langle E_1^0|\mathbf{x}'\rangle g(x^2 + y^2 + z^2)\delta^3(\mathbf{x} - \mathbf{x}')\langle \mathbf{x}|E_1^0\rangle \\
&= \int_{-\infty}^{\infty} \mathrm{d}^3x\langle E_1^0|\mathbf{x}\rangle gr^2\langle \mathbf{x}|E_1^0\rangle = \frac{g}{\pi r_0^3}\int_0^{\infty}\mathrm{d}r\ r^2\int_0^{\pi}\mathrm{d}\theta\ \sin\theta\int_0^{2\pi}\mathrm{d}\phi\ r^2e^{-2r/r_0} \\
&= \frac{g}{\pi r_0^3}(4\pi)\int_0^{\infty}\mathrm{d}r\ r^4e^{-2r/r_0} = 3gr_0^2\,.
\end{aligned}$$

As a result of the perturbation, the ground state energy is raised an amount $3gr_B^2$.

Example 4.2.3 Use first-order perturbation theory to calculate the $n = 2$, $\ell = 1$ energy levels of the hydrogen atom in the presence of the perturbation $H_1 = g(Q_x^2 + Q_y^2 + Q_z^2)$ of the previous example. The $n = 2$, $\ell = 1$ hydrogen atom wave functions are $\langle \mathbf{x}|E_2, \ell = 1, m = 0\rangle = R_{n=2,\ell=1}(r)Y_{\ell=1}^m$ The radial wave function and spherical harmonics are tabulated, respectively, in Tables 3.1 and 3.2.

Solution As discussed in the previous example, the eigenstates of both the free Hamiltonian and the "exact" Hamiltonian can be labeled by the quantum numbers

$b = \ell, m$. Therefore, although the $n = 2$; $\ell = 1$; $m = \pm 1, 0$ energy levels are degenerate, the eigenvector $|E^{(0)}_{n=2}, \ell = 1, m)$ of H_0 is the best approximation to the eigenvector $|E_{n=2}, \ell = 1, m\rangle$ of H. Consequently it is possible to calculate the energy levels of the "exact" Hamiltonian to first order in the perturbation using the formulas that have just been derived. Putting the quantum numbers $b = \ell, m$ back into (4.2.20),

$$E^{(1)}_{n=2,\ell=1,m} = E^0_{n=2} + \langle E^0_{n=2,\ell=1,m}|H_1|E^0_{n=2,\ell=1,m}\rangle \,.$$

Rewriting the matrix element as an integral over wave functions as was done in the previous example and using the explicit expressions for hydrogen atom wave functions,

$$\begin{aligned} E^{(1)}_{n=2,\ell=1,m} &= E^0_{n=2} \\ &+ g\int_0^\infty \mathrm{d}r\, r^2 \left(\frac{1}{\sqrt{3}(2r_B)^{3/2}}\frac{r}{r_0}e^{-r/2r_0}\right)^2 r^2 \int_0^{2\pi}\mathrm{d}\phi \int_0^\pi \mathrm{d}\theta\ \sin\theta |Y^{\ell=1}_m(\theta,\phi)|^2 \,. \end{aligned}$$

As can be readily checked, the integral over the spherical harmonic is unity. Evaluating the radial integral,

$$E^{(1)}_{n=2,\ell=1,m} = E^0_{n=2} + 30\, g r_0^2 .$$

Example 4.2.4 A rotator in a uniform magnetic field B_3 experiences a perturbation $\alpha(J_2)^2$ where α is a constant. Using first-order perturbation theory calculate the energy eigenvalues and eigenvectors.

Solution From (4.1.14), the Hamiltonian H_0 of a rotator in a uniform magnetic field pointing in the 3-direction is

$$H_0 = \frac{1}{2I}\mathbf{J}^2 - g\frac{q}{2m}B_3 J_3 \quad \text{and} \quad H_1 = \alpha(J_2)^2 .$$

For this problem a complete set of commuting operators for the free Hamiltonian consists of H_0 and J_3. But J_3 does not commute with H, and a complete set of commuting operators for the "exact" Hamiltonian consists of H and $-g(q/2m)B_3 J_3 + \alpha(J_2)^2$. In spite of the fact that an operator B that commutes with both H_0 and H does not exist, it is still possible to use (4.2.20) and (4.2.22), respectively, to calculate the eigenvalues and eigenvectors of H to first order in α because the eigenvalues of H_0 are non-degenerate. For this problem

$$|E^0_n\rangle = |E^0_{j,j_3}, j_3\rangle \equiv |j, j_3\rangle \ \text{ and } \ E^0_{j,j_3} = \frac{\hbar^2 j(j+1)}{2I} - g\frac{q\hbar}{2m}B_3 j_3 \,,$$

where $|j, j_3\rangle$ is the usual angular momentum basis (2.3.35). To calculate the first-order correction to E^0_{j,j_3} for this problem, (4.2.20) becomes

$$E^{(1)} = E^0_{j,j_3} + \langle j, j_3|\alpha(J_2)^2|j, j_3\rangle\ .$$

Using $J_2 = -(i/2)(J_+ - J_-)$ to rewrite $(J_2)^2$ in terms of the angular momentum raising and lowering operators $J_\pm$,

$$\begin{aligned} E^{(1)} &= E^0_{j,j_3} + \alpha\langle j, j_3|\left[-\frac{i}{2}(J_+ - J_-)\right]^2|j, j_3\rangle\,, \\ &= E^0_{j,j_3} - \frac{\alpha}{4}\langle j, j_3|J_+^2 + J_-^2 - J_+J_- - J_-J_+|j, j_3\rangle\,. \end{aligned}$$

Only the last two terms have non-zero diagonal matrix elements. Using (2.3.43b) and (2.3.43c),

$$\begin{aligned} E^{(1)} &= E^0_{j,j_3} + \frac{\alpha}{4}\hbar^2\left[(j + j_3)(j - j_3 + 1) + (j - j_3)(j + j_3 + 1)\right] \\ &= E^0_{j,j_3} + \frac{\alpha}{2}\hbar^2\left[(j + j_3)(j - j_3) + j\right]\,. \end{aligned}$$

The eigenvectors of H are calculated to first order in α by employing (4.2.22),

$$|E^{(1)}\rangle = |j, j_3\rangle + \sum_{j_3'\neq j_3}|j, j_3\rangle\frac{\langle j, j_3'|\alpha(J_2^2)|j, j_3\rangle}{E^0_{j,j_3} - E^0_{j,j_3'}}\,.$$

On the right-hand side there is no sum over j since $\mathbf{J}^2$ commutes with both H_0 and H, and $|E^{(1)})$ is an eigenstate of $\mathbf{J}^2$ with angular momentum j. Rewriting $(J_2)^2$ in terms of raising and lowering operators,

$$|E^{(1)}\rangle = |j, j_3\rangle + -\frac{\alpha}{4}\sum_{j_3'\neq j_3}|j, j_3'\rangle\frac{\langle j, j_3'|J_+^2 + J_-^2 - J_+J_- - J_-J_+|j, j_3\rangle}{E^0_{j,j_3} - E^0_{j,j_3'}}\,.$$

The operator $(J_+)^2$ has a non-zero matrix element only when $j_{3'} = j_3 + 2$, the operator $(J_-)^2$ has a non-zero matrix element only when $j_{3'} = j_3 - 2$, and the operators J_+J_- and J_-J_+ do not contribute since the term with $j_3 = j_{3'}$ is not included in the sum. Therefore,

$$\begin{aligned} |E^{(1)}\rangle = |j, j_3\rangle &- \frac{\alpha}{4}\frac{\langle j, j_3 + 2|J_+^2|j, j_3\rangle}{E^0_{j,j_3} - E^0_{j,j_3+2}}|j, j_3 + 2\rangle \\ &- \frac{\alpha}{4}\frac{\langle j, j_3 - 2|J_-^2|j, j_3\rangle}{E^0_{j,j_3} - E^0_{j,j_3-2}}|j, j_3 - 2\rangle\,. \end{aligned}$$

Evaluating the matrix elements and using the explicit expression for E^0_{j,j_3},

$$|E^{(1)}\rangle = |j, j_3\rangle - \frac{\alpha\hbar m}{4gqB_3}\sqrt{(j-j_3)(j-j_3+1)(j+j_3+1)\langle j+j_3+2)}\,|j, j_3+2\rangle$$
$$+\frac{\alpha\hbar m}{4gqB_3}\sqrt{(j+j_3)(j+j_3-1)(j-j_3+1)(j-j_3+2)}\,|j, j_3-2\rangle\,.$$

In this example the quantum numbers that label the eigenvectors $|E^{(1)}\rangle$ of H have not been explicitly listed. Since H and $-g(q/2m)B_3J_3 + \alpha(J_2)^2$ constitute a complete set of commuting operators, the eigenvalues of these two operators could be used to completely specify the eigenvectors of H.

Second-Order Approximation

To obtain the second-order approximation, the first-order approximations for E_n and $|E_n\rangle$ are substituted into the right-hand sides of (4.2.11) and (4.2.17). The results are as follows (Problems 4.14 and 4.15):

$$E^{(2)}_n = E^0_n + \langle E^0_n|H_1|E^0_n\rangle + \sum_{n'\neq n}\frac{|\langle E^0_n|H_1|E^0_{n'}\rangle|^2}{E^0_n - E^0_{n'}}\,, \tag{4.2.23}$$

and

$$|E^{(2)}_n\rangle = |E^0_n\rangle + \sum_{n'\neq n}|E^0_{n'}\rangle\frac{\langle E^0_{n'}|H_1|E^0_n\rangle^2}{E^{(1)}_n - E^0_{n'}}$$
$$+\sum_{n'\neq n}\sum_{n''\neq n}|E^0_{n'}\rangle\frac{\langle E^0_{n'}|H_1|E^0_{n''}\rangle}{E^{(1)}_n - E^0_{n'}}\,\frac{\langle E^0_{n''}|H_1|E^0_n\rangle}{E^0_n - E^0_{n''}}\,. \tag{4.2.24}$$

Continuing in this fashion it is possible to obtain higher-order approximations for the exact energy eigenvalues and eigenvectors.

Note that the eigenvectors $|E^{(1)}_n)$ are neither normalized nor orthogonal. However, since the matrix element of H_1 is small, from (4.2.22) it follows that the deviations from orthogonality are small. Similarly, the eigenvectors $|E^{(2)}_n\rangle$ are also neither normalized nor orthogonal. The exact eigenvectors $|E_{n,b}, b\rangle \equiv |E_n\rangle$ in (4.2.17) could be normalized to unity. Instead it is conventional to choose the normalization

$$\langle E^0_n|E_n\rangle = 1\,, \tag{4.2.25}$$

which also fixes the relative phase of the “exact” eigenvector $|E_n\rangle$ and the free eigenvector $|E^0_n\rangle$. With the above choice for normalization, the exact Eqs. (4.2.11) and (4.2.17) become, respectively,

$$E_n = E^0_n + \langle E^0_n|H_1|E_n\rangle\,, \tag{4.2.26}$$

and

$$|E_n\rangle = |E_n^0\rangle + \sum_{n' \neq n} |E_{n'}^0\rangle \frac{\langle E_{n'}^0|H_1|E_n\rangle}{E_n - E_{n'}^0} \, . \tag{4.2.27}$$

If the iteration procedure for solving (4.2.11) and (4.2.17) is to make sense, the exact eigenvector $|E_n\rangle$ must not differ much from $|E_n^0\rangle$. From (4.2.17) it is possible to conclude that

$$|\langle E_{n'}^0|H_1|E_n\rangle| \ll |E_n - E_{n'}^0| \, . \tag{4.2.28}$$

But even if the above condition is satisfied, it does not always ensure that the lowest order solution is actually an approximation to the exact solution. However, there is a large class of problems for which the approximation does work, including the spectra of atoms and molecules where H_0 is given by the Coulomb interaction. The interaction Hamiltonian H_1 could describe the interaction with a weak electromagnetic field or the interaction between the spin of the electrons and the magnetic field due to nuclear motion.

Example 4.2.5 Consider the Hamiltonian $H = H_0 + H_1$ in one-dimensional space where $H_0 = P^2/2\mu + (1/2)kQ^2$ and $H_1 = \lambda Q$.

(a) Calculate the exact energy eigenvalues E_n.
(b) Express the exact energy eigenvectors $|E_n\rangle$ in terms of operators and the exact ground-state energy eigenvector $|E_0\rangle$.
(c) Use second-order perturbation theory to calculate the energy $E_n^{(2)}$ and verify that E_n and $E_n^{(2)}$ agree to second order in λ.
(d) Calculate the energy eigenvectors $|E_n^{(1)}\rangle$ using first-order perturbation theory.
(e) In the formula in (b), substitute the expression for $|E_0^{(1)}\rangle$ calculated from perturbation theory and verify that $|E_n^{(1)}\rangle$ calculated from (b) agrees with the expression calculated in (d) to first order in λ.

Solution

(a) As can readily be checked, by "completing the square" the Hamiltonian can be written in the form

$$H = \frac{P^2}{2\mu} + \frac{1}{2}k(Q + \frac{\lambda}{k})^2 - \frac{\lambda^2}{2k} \, , = \frac{P^2}{2\mu} + \frac{1}{2}k(Q')^2 - \frac{\lambda^2}{2k} \, .$$

Since P and Q' obey the Heisenberg commutation relations (1.2.11), it immediately follows from (1.2.40) that the exact energy eigenvalues are

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) - \frac{\lambda^2}{2k}, \quad n = 0, 1, \ldots \, ,$$

(b) From (1.2.49)

$$|E_n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|E_0\rangle\,,$$

where

$$a^\dagger = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\,(Q+\frac{\lambda}{k}) - \frac{i}{\sqrt{\mu\omega\hbar}}\,P\right)\,.$$

To calculate $E_n^{(2)}$ and $|E_n^{(1)})$ from perturbation theory, the relations

$$E_n^0 = \hbar\omega\left(n+\frac{1}{2}\right)\,,$$

and

$$\langle E_n^0|Q|E_{n'}^0\rangle = \sqrt{\frac{\hbar}{2\mu\omega}}\left[\sqrt{n'+1}\;\delta_{n,n'+1} + \sqrt{n'}\;\delta_{n,n'-1}\right]\,,$$

given, respectively, by (1.2.40) and (1.2.52) are needed.

(c) Using (4.2.24) and noting that the only non-zero matrix elements $\langle E_n^0|Q|E_{n'}^0\rangle$ in the sum occur for $n'=n-1$ and $n'=n+1$,

$$\begin{aligned}E_n^{(2)} &= E_n^0 + \langle E_n^0|\lambda Q|E_n^0\rangle + \frac{|\langle E_n^0|\lambda Q|E_{n-1}^0\rangle|^2}{E_n^0-E_{n-1}^0} + \frac{|\langle E_n^0|\lambda Q|E_{n+1}^0\rangle|^2}{E_n^0-E_{n+1}^0}\\ &= \hbar\omega\left(n+\frac{1}{2}\right) + 0 + \frac{\left(\lambda\sqrt{\frac{\hbar}{2\mu\omega}}\sqrt{n}\right)^2}{\hbar\omega\left(n+\frac{1}{2}\right)-\hbar\omega\left(n-1+\frac{1}{2}\right)}\\ &\quad + \frac{\left(\lambda\sqrt{\frac{\hbar}{2\mu\omega}}\sqrt{n+1}\right)^2}{\hbar\omega\left(n+\frac{1}{2}\right)-\hbar\omega\left(n+1+\frac{1}{2}\right)} = \hbar\omega\left(n+\frac{1}{2}\right) - \frac{\lambda^2}{2k}.\end{aligned}$$

The above equation agrees with the exact expression E_n to order λ^2. In fact, for this specific perturbation $E_n^{(2)}$ and E_n agree exactly.

(d) Noting that $\langle E_n^0|Q|E_{n'}^0\rangle$ is non-zero only when $n'=n-1$ and $n'=n+1$, Eq. (4.2.22) for $|E_n^{(1)}\rangle$ becomes

$$|E_n^{(1)}\rangle = |E_n^0\rangle + |E_{n-1}^0\rangle\frac{\langle E_{n-1}^0|\lambda Q|E_n^0\rangle}{E_n^0-E_{n-1}^0} + |E_{n+1}^0\rangle\frac{\langle E_{n+1}^0|\lambda Q|E_n^0\rangle}{E_n^0-E_{n+1}^0}\,.$$

Employing the explicit expression for $\langle E_n^0|Q|E_{n'}^0\rangle$ and E_i^0,

$$|E_n^{(1)}\rangle = |E_n^0\rangle + \frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}\sqrt{n}|E_{n-1}^0\rangle - \frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}\sqrt{n+1}|E_{n+1}^0\rangle \, .$$

(e) Making the replacement $E_n \to E_n^{(1)}$ in (b),

$$|E_n^{(1)}\rangle \simeq \frac{1}{\sqrt{n!}}\left[\frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\,(Q+\frac{\lambda}{k}) - \frac{i}{\sqrt{\mu\omega\hbar}}\,P\right)\right]^n |E_0^{(1)}\rangle \, .$$

Using (d) to express $|E_0^{(1)}\rangle$ in terms of $|E_0^{(0)}\rangle$ and $|E_1^{(0)}\rangle$ and defining the "free" raising operator

$$a^{(0)\dagger} = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\,Q - \frac{i}{\sqrt{\mu\omega\hbar}}\,P\right) ,$$

$$|E_n^{(1)}\rangle \simeq \frac{1}{\sqrt{n!}}\left(a^{(0)\dagger} + \frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}\right)^n \left[|E_0^{(0)}\rangle - \frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}|E_1^{(0)}\rangle\right] .$$

Keeping only the first two terms in the binomial expansion,

$$|E_n^{(1)}\rangle \simeq \frac{1}{\sqrt{n!}}\left[(a^{(0)\dagger})^n + \frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}(a^{(0)\dagger})^{n-1} + \cdots\right]\left[|E_0^{(0)}\rangle - \frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}|E_1^{(0)}\rangle\right] .$$

Deleting the term proportional to λ^2,

$$\begin{aligned}|E_n^{(1)}\rangle = \frac{1}{\sqrt{n!}}(a^{(0)\dagger})^n|E_0^{(0)}\rangle + \frac{n}{\sqrt{n!}}\frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}(a^{(0)\dagger})^{n-1}|E_0^{(0)}\rangle \\ - \frac{1}{\sqrt{n!}}\frac{\lambda}{k}\sqrt{\frac{\mu\omega}{2\hbar}}(a^{(0)\dagger})^n|E_1^{(0)}\rangle \, .\end{aligned}$$

Using

$$(a^{(0)\dagger})^m\,|E_n^{(0)}\rangle = \sqrt{\frac{(n+m)!}{n!}}|E_{n+m}^{(0)}\rangle \, ,$$

that immediately follows from (1.2.50), the final equation in (d) is obtained, implying that $|E_n^{(1)}\rangle$ calculated in (d) and (e) agree as they must.

In summary, the formulas derived in this section express the eigenvalues and eigenvector of the "exact" Hamiltonian H in terms of the eigenvalues and eigenvectors of the free Hamiltonian H_0. These formulas can be used if the energy eigenvalues of H_0 are non-degenerate or if it is possible to find a set of hermitian operators B that commute with both H_0 and H such that H_0 and B as well as H and B form complete sets of commuting observables.

4.3 Degenerate, Stationary Perturbation Expansion

In the previous section the following two cases were considered:

Case I A hermitian operator B is known that commutes with both H_0 and H_1 where B and H_0 form a complete set of commuting observables.

Case II The eigenvalues of H_0 are non-degenerate.

For either of these cases, the zeroth-order approximation to each of the eigenvectors of the "exact" Hamiltonian can be immediately identified and is a single eigenvector of H_0.

In this section the remaining case is considered:

Case III The energy eigenvalues of the free Hamiltonian H_0 are degenerate, and no hermitian operator B is known that commutes with both H_0 and H.

For the case now being considered, H_0 and B form a complete set of commuting observables, but B does not commute with the "exact" Hamiltonian H. A complete set of commuting operators for the "exact" Hamiltonian consists of H and operators D with eigenvalues d. The operators D are assumed not to commute with H_0 since the situation would then reduce to Case I considered in the previous section. As in the previous section, it is still assumed that to each eigenvalue $E_{n,d}$ of H there corresponds an eigenvalue E_n^0 of H_0 such that

$$\begin{pmatrix} E_{n,d_1} \\ E_{n,d_2} \\ \vdots \\ E_{n,d_m} \end{pmatrix} \rightarrow E_n^0 \quad \text{as} \quad H_1 \rightarrow 0\,. \tag{4.3.1}$$

As a consequence, m eigenvalues of H correspond to one eigenvalue of H_0, which is m-degenerate where m is an integer or infinity.

It is not difficult to understand how perturbation theory must be modified for the case being considered. First attention is restricted to the simplest situation where two eigenstates of H_0, denoted $|E_{n,b_1}^0\rangle$ and $|E_{n,b_2}^0\rangle$, have the same energy eigenvalue E_n^0. When a perturbation H_1 is applied, two new stationary states are obtained, $|E_{n,d_1}, d_1\rangle$ and $|E_{n,d_1}, d_2\rangle$ that are eigenstates of the exact Hamiltonian $H = H_0 + H_1$ with respective energy eigenvalues E_{n,d_1} and E_{n,d_2}. When the perturbation is "turned off", it is *not* necessarily true that $|E_{n,d_1}, d_1\rangle \rightarrow |E_{n,b_1}^0\rangle$ and $|E_{n,d_1}, d_2\rangle \rightarrow |E_{n,b_2}^0\rangle$. Instead, in general, as the interaction is "turned off", each of the exact eigenvectors approaches some linear combination of the unperturbed eigenvectors:

$$\lim_{H_1\to 0} |E_{n,d_1}, d_1\rangle = \alpha_1|E_n^0, b_1\rangle + \alpha_2|E_n^0, b_2\rangle\,, \tag{4.3.2a}$$

$$\lim_{H_1\to 0} |E_{n,d_2}, d_2\rangle = \alpha_3|E_n^0, b_1\rangle + \alpha_4|E_n^0, b_2\rangle\,. \tag{4.3.2b}$$

In the above two equations each of the α_i is a constant. As indicated in (4.3.1), when the energy eigenvalues are degenerate, both of the exact energy eigenvalues E_{n,d_1} and E_{n,d_1} approach E_n^0 as the perturbation H_1 is "turned off". That is, the energy eigenvalue that is almost correct is known. But as the perturbation is turned off, the constants α_i in (4.3.2) are not known so the energy eigenvector of H_0 that is the best approximation to each energy eigenvector of H is not known.

Since the energy eigenvector of H_0 that is the best approximation to each energy eigenvector of H must be determined, it is written in the form

$$\lim_{H_1\to 0} |E_{n,d}, d\rangle \equiv |E_n^0, d\rangle = c_1|E_n^0, b_1\rangle + c_2|E_n^0, b_2\rangle\,, \tag{4.3.3}$$

as required by (4.3.2). In the above equation, d represents either of the eigenvalues d_1 or d_2 just as j_3 represents either of the eigenvalues 1/2 or $-1/2$ when $j = 1/2$. The constants c_1 and c_2, which are unknown and must be determined, take on different values for $d = d_1$ and $d = d_2$. In the following calculations $|E_{n,b_1}^0\rangle$ and $|E_{n,b_2}^0\rangle$ are assumed to be orthonormal,

$$\langle E_n^0, b_i|E_n^0, b_j\rangle = \delta_{ij}\,. \tag{4.3.4}$$

The two degenerate eigenvectors $|E_{n,b_i}^0\rangle$ are orthogonal because $b_i \neq b_j$, and they can always be normalized so that (4.3.4) is satisfied.

The energy eigenvectors on the right-hand side of (4.3.2) are called *stabilized energy eigenvectors*. The eigenvector (4.3.3) is also a stabilized energy eigenvector, and, as will be seen shortly, there are two solutions for c_1 and c_2: the two solutions on the right-hand side of (4.3.2). The stabilized energy eigenvectors are the "correct" energy eigenvectors to zeroth order in the perturbation.

In the following discussion of degenerate perturbation theory, the stabilized energy eigenvectors and the first-order corrections to the energy will be determined. The starting point is the exact energy eigenvalue equation,

$$H|E_{n,d}, d\rangle = (H_0 + H_1)|E_{n,d}, d\rangle\,, \tag{4.3.5}$$

which implies

$$E_{n,d}|E_{n,d}^0, d\rangle = (H_0 + H_1)|E_{n,d}, d\rangle\,. \tag{4.3.6}$$

Taking the scalar product of the above equation with one of the degenerate eigenvectors $|E_n^0, b_i\rangle$, where $i = 1$ or 2,

$$E_{n,d}\langle E_n^0, b_i|E_{n,d}, d\rangle = \langle E_n^0, b_i|H_0 + H_1|E_{n,d}, d\rangle. \tag{4.3.7}$$

Using the fact that H_0 is hermitian and that

$$H^{(0)}|E_n^0, b_i\rangle = E_n^0|E_n^0, b_i\rangle\,, \tag{4.3.8}$$

where E_n^0 is the degenerate energy eigenvalue, (4.3.7) can be written in the desired form,

$$(E_{n,d} - E_n^0)\langle E_n^0, b_i|E_{n,d}, d\rangle = \langle E_n^0, b_i|H_1|E_{n,d}, d\rangle\,. \tag{4.3.9}$$

When energy eigenvalues are calculated only to first order in the perturbation and the eigenvectors only to zeroth order, (4.3.9), which is essentially the same as (4.2.10), is the fundamental equation in degenerate, stationary perturbation theory.

Zeroth-Order Approximation
The zeroth-order approximation is obtained by setting $H_1 = 0$. Thus,

$$E_{n,d} \to E_n^0 \tag{4.3.10}$$

and, as indicated in (4.3.3), the "exact" eigenvector $|E_{n,d}, d\rangle$ is a linear combination of the unperturbed, degenerate eigenvectors

$$|E_{n,d}, d\rangle \to c_1|E_n^0, b_1\rangle + c_2|E_n^0, b_2\rangle \equiv |E_n^0, d\rangle\,. \tag{4.3.11}$$

The constants c_1 and c_2 remain unknown for the moment but will be determined by the first-order approximation.

First-Order Approximation
The energy will now be determined to first-order in the perturbation and, by calculating the constants c_1 and c_2 in the above equation, the energy eigenvectors will be determined to zeroth-order. The first-order approximation $E_{n,d}^{(1)}$ for energy is obtained by replacing the exact energy eigenvector $|E_{n,d}, d\rangle$ in (4.3.9) by its zeroth-order approximation $|E_n^0, d\rangle$ as given in (4.3.11). This is the same approximation made to obtain the first-order approximation for energy in the previous section. Making the indicated substitution, (4.3.9) becomes

$$(E_{n,d}^{(1)} - E_n^0)\langle E_n^0, b_i|E_n^0, d\rangle = \langle E_n^0, b_i|H_1|E_n^0, d\rangle\,. \tag{4.3.12}$$

The above equation corresponds to the first-order approximation (4.2.19) in the previous section. Taking the index $i = 1$ in the above equation, using the explicit expression for $|E_n^0, d\rangle$ as given in (4.3.11), and recalling from (4.3.4) that the unperturbed energy eigenvectors are orthonormal, (4.3.12) becomes

$$(E_{n,d}^{(1)} - E_n^0)c_1 = \langle E_n^0, b_1|H_1|[c_1|E_n^0, b_1\rangle + c_2|E_n^0, b_2\rangle] \,. \tag{4.3.13a}$$

Similarly, taking the index $i = 2$ in (4.3.12),

$$(E_{n,d}^{(1)} - E_n^0)c_2 = \langle E_n^0, b_2|H_1|[c_1|E_n^0, b_1\rangle + c_2|E_n^0, b_2\rangle] \,. \tag{4.3.13b}$$

Introducing the notation

$$\langle E_n^0, b_i|H_1|E_n^0, b_j\rangle \equiv \langle H_1\rangle_{ij} \,, \tag{4.3.14}$$

Equation (4.3.13) can be written in the desired form:

$$(E_{n,d}^{(1)} - E_n^0)c_1 = c_1\langle H_1\rangle_{11} + c_2\langle H_1\rangle_{12} \,, \tag{4.3.15a}$$

$$(E_{n,d}^{(1)} - E_n^0)c_2 = c_1\langle H_1\rangle_{21} + c_2\langle H_1\rangle_{22} \,. \tag{4.3.15b}$$

Writing the two equations in matrix form,

$$\begin{pmatrix} E_{n,d}^{(1)} - E_n^0 - \langle H_1\rangle_{11} & -\langle H_1\rangle_{12} \\ -\langle H_1\rangle_{21} & E_{n,d}^{(1)} - E_n^0 - \langle H_1\rangle_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0 \,. \tag{4.3.16}$$

The matrix equation has a non-trivial solution ($c_1 = c_2 = 0$ is the trivial solution) only if the determinant of the above matrix vanishes,

$$\begin{vmatrix} E_{n,d}^{(1)} - E_n^0 - \langle H_1\rangle_{11} & -\langle H_1\rangle_{12} \\ -\langle H_1\rangle_{21} & E_{n,d}^{(1)} - E_n^0 - \langle H_1\rangle_{22} \end{vmatrix} = 0 \,. \tag{4.3.17}$$

Note that $H_1 = H_1^\dagger$, which implies

$$\langle H_1\rangle_{12} = \langle H_1\rangle_{21}^* \,. \tag{4.3.18}$$

Evaluating the determinant and then using the quadratic formula, the first-order approximation for energy is obtained (Problem 4.15):

$$E_{n,d}^{(1)} = E_n^0 + \frac{1}{2}\left(\langle H_1\rangle_{11} + \langle H_1\rangle_{22} \pm \sqrt{(\langle H_1\rangle_{11} - \langle H_1\rangle_{22})^2 + 4|\langle H_1\rangle_{12}|}\right)^2 \,. \tag{4.3.19}$$

From the above formula, it follows that there are two solutions for the first-order approximation $E_{n,d}^{(1)}$ unless $\langle H_1\rangle_{11} = \langle H_1\rangle_{22}$ and $\langle H_1\rangle_{12} = 0$. When a small perturbation is applied, the energy levels of the two unperturbed degenerate states are shifted by different amounts and the degeneracy is removed except when the above two equalities are satisfied.

Substituting the expression for $E_{n,d}^{(1)}$ as given in (4.3.19) into (4.3.16) (Problem 4.15),

$$2\langle H_1\rangle_{21}c_1 = \left(\langle H_1\rangle_{11} - \langle H_1\rangle_{22} \pm \sqrt{(\langle H_1\rangle_{11} - \langle H_1\rangle_{22})^2 + 4|\langle H_1\rangle_{12}|^2}\right)c_2\,. \tag{4.3.20}$$

If the matrix element $\langle H_1\rangle_{12} \neq 0$, the two unnormalized, stabilized energy eigenvectors corresponding, respectively, to the two energies in (4.3.19) are

$$\begin{aligned}|E_n^0, d\rangle &= \frac{c_2}{2\langle H_1\rangle_{21}}\left(\langle H_1\rangle_{11} - \langle H_1\rangle_{22} \pm \sqrt{(\langle H_1\rangle_{11} - \langle H_1\rangle_{22})^2 + 4|\langle H_1\rangle_{12}|^2}\right)|E_n^0, b_1\rangle \\ &\quad + c_2|E_n^0, b_2\rangle\,. \end{aligned} \tag{4.3.21}$$

As can be readily verified, the above two stabilized energy eigenvectors are orthogonal (Problem 4.16).

If the matrix element $\langle H_1\rangle_{12} = 0$, which also implies $\langle H_1\rangle_{21} = 0$, then the two equations in (4.3.16) become,

$$(E_{n,d}^{(1)} - E_n^0 - \langle H_1\rangle_{11})c_1 = 0\,, \tag{4.3.22a}$$

$$(E_{n,d}^{(1)} - E_n^0 - \langle H_1\rangle_{22})c_2 = 0\,. \tag{4.3.22b}$$

There are two solutions to (4.3.22): The first solution is obtained by requiring that the term in parenthesis in (4.3.22a) and c_2 in (4.3.22b) be zero. The second solution is obtained by requiring that c_1 in (4.3.22a) and the term in parenthesis in (4.3.22b) be zero. The energies and the corresponding stabilized energy eigenvectors are

Energy	Stabilized Energy Eigenvector	
$E_{n,d}^{(1)} = E_n^0 + \langle H_1\rangle_{11}$	$\lvert E_n^0, b_1\rangle$	(4.3.23a)
$E_{n,d}^{(1)} = E_n^0 + \langle H_1\rangle_{22}$	$\lvert E_n^0, b_2\rangle$	(4.3.23b)

When the matrix element $\langle H_1\rangle_{12} = 0$, the original degenerate vectors $|E_n^0, b_1\rangle$ and $|E_n^0, b_2\rangle$ are the stabilized energy eigenvectors, and the first-order corrections to energy are identical to those obtained from the formula (4.2.20).

Note that if $\langle H_1\rangle_{12} = 0$ *and* $\langle H_1\rangle_{11} = \langle H_1\rangle_{22}$, the stabilized energy eigenvectors are degenerate to first-order in the perturbation. The constants c_1 and c_2 in (4.3.22) are then arbitrary, so any linear combination of $|E_n^0, b_1\rangle$ and $|E_n^0, b_2\rangle$ is a stabilized energy eigenvector.

Example 4.3.1 Consider the Hamiltonian $H = H_0 + H_1$ where

$$H_0 = \frac{\mathbf{S}^2}{2I}\,, \quad H_1 = -g\frac{q}{2m}B_1 S_1\ .$$

Determine the stabilized energy eigenvectors and the corresponding energies to first order in the perturbation H_1 for spin $s = 1/2$.

Solution This problem can be solved exactly by using the energy eigenstates $|s = 1/2,\ s_1 = \pm 1/2\rangle$, but it will first be solved using the basis $|s = 1/2,\ s_3 = \pm 1/2\rangle$ to illustrate how the formalism of degenerate perturbation theory is applied.

The vectors $|s = 1/2,\ s_3 = 1/2\rangle$ and $|s = 1/2,\ s_3 = -1/2\rangle$ are both eigenvectors of H_0 with an energy eigenvalue $\hbar^2 s(s+1)/2I = 3\hbar^2/8I$. The stabilized energy eigenvector $|E^0\rangle$ is chosen to be of the form

$$|E^0\rangle = c_1|1/2\ ,\ 1/2\rangle + c_2|1/2\ ,\ -1/2\rangle\,.$$

To calculate the required matrix elements of H_1, note that

$$J_1 = \frac{1}{2}\left[J_1 + iJ_2\right] + \frac{1}{2}\left[J_1 - iJ_2\right] = \frac{1}{2}\left(J_+ + J_-\right)\ .$$

Then, for example,

$$\begin{aligned}\langle H_1\rangle_{12} &= \langle 1/2, 1/2| - g\frac{q}{2m}B_1 J_1|1/2, -1/2\rangle\,,\\ &= -g\frac{q}{2m}B_1\langle 1/2, 1/2|\frac{1}{2}(J_+ + J_-)|1/2, -1/2\rangle\,.\end{aligned}$$

With the aid of (2.3.43) ,

$$\langle H_1\rangle_{12} = -\frac{gq\hbar}{4m}B_1\ .$$

In a similar fashion,

$$\langle H_1\rangle_{11} = \langle H_1\rangle_{22} = 0, \quad \langle H_1\rangle_{21} = \langle H_1\rangle_{12}^* = -\frac{gq\hbar}{4m}B_1\ .$$

Using the above matrix elements, (4.3.16) becomes

$$\begin{pmatrix} E^{(1)}_{s,s_1} - E^0_s & \frac{gq\hbar}{4m}B_1 \\ \frac{gq\hbar}{4m}B_1 & E^{(1)}_{s,s_1} - E^0_s \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0\,.$$

The energy eigenvalues are found by setting the determinant of the above matrix to zero,

$$E_{s,s_1}^{(1)} = E_s^0 \pm \frac{gq\hbar}{4m} B_1 = \frac{3\hbar^2}{8I} \pm \frac{gq\hbar}{4m} B_1 \,.$$

To obtain the stabilized energy eigenvectors, the above values for $E_{s,s_1}^{(1)}$ are substituted into the matrix equation yielding

$$\pm\frac{gq\hbar}{4m} B_1\, c_1 + \frac{gq\hbar}{4m} B_1\, c_2 = 0\,, \qquad \text{or} \qquad c_2 = \mp c_1\,.$$

Taking the top sign and then the bottom sign, the energy eigenvalues and the corresponding stabilized energy eigenvectors are as follows:

Energy	Stabilized Energy Eigenvector
$E_{s,s_1}^{(1)} = \frac{3\hbar^2}{8I} + \frac{gq\hbar}{4m} B_1$	$\frac{1}{\sqrt{2}}[\|1/2, 1/2\rangle - \|1/2, -1/2\rangle]$
$E_{s,s_1}^{(1)} = \frac{3\hbar^2}{8I} - \frac{gq\hbar}{4m} B_1$	$\frac{1}{\sqrt{2}}[\|1/2, 1/2\rangle + \|1/2, -1/2\rangle]$

As can be readily checked (Problem 4.17), the stabilized energy eigenvectors are the eigenvectors $|s = 1/2, s_1 = \pm 1/2\rangle$ that satisfy

$$S_1 |s = 1/2, s_1 = \pm 1/2\rangle = \pm\hbar(1/2)|s = 1/2, s_1 = \pm 1/2\rangle.$$

That is,

$$|s = 1/2, s_1 = \pm 1/2\rangle = \frac{1}{\sqrt{2}}[|1/2, 1/2\rangle \pm |1/2, -1/2\rangle]\,.$$

The formulas for energy that have been derived to first order in H_1 are actually exact as can be immediately verified using the basis $|s, s_1\rangle$.

In the previous example (4.3.19) and (4.3.21) could have been used, respectively, to determine the energy eigenvalues $E_n^{(1)}$ and the stabilized energy eigenvectors. Instead (4.3.16) was employed, which incorporates both (4.3.19) and (4.3.21), is easy to remember, and is readily generalized to the case where there are N degenerate eigenvectors of H_0.

If there are N degenerate eigenvectors of H_0, a stabilized energy eigenvector $|E_n^0, d\rangle$ is of the form

$$|E_n^0, d\rangle = \sum_{j=1}^{N} c_j |E_n^0, b_j\rangle. \tag{4.3.24}$$

Substituting the above equation into (4.3.9),

$$(E^{(1)}_{n,d} - E^0_n)\langle E^0_n, b_i| \left[\sum_{j=1}^{N} c_j |E^0_n, b_j\rangle\right] = \langle E^0_n, b_i|H_1| \left[\sum_{j=1}^{N} c_j |E^0_n, b_j\rangle\right] . \tag{4.3.25}$$

By taking $j = 1, 2, \ldots, N$ in (4.3.25), N equations are obtained that can be written as the matrix

$$\begin{pmatrix} E^{(1)}_{n,d} - E^0_n - \langle H_1\rangle_{11} & -\langle H_1\rangle_{12} & -\langle H_1\rangle_{13} & \cdots \\ -\langle H_1\rangle_{21} & E^{(1)}_{n,d} - E^0_n - \langle H_1\rangle_{22} & -\langle H_1\rangle_{23} & \cdots \\ -\langle H_1\rangle_{31} & -\langle H_1\rangle_{32} & E^{(1)}_{n,d} - E^0_n - \langle H_1\rangle_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = 0 . \tag{4.3.26}$$

Example 4.3.2 The Hamiltonian for a harmonic oscillator in two-dimensional space is

$$H_0 = \frac{P_x^2 + P_y^2}{2m} + \frac{1}{2}k(Q_x^2 + Q_y^2) .$$

Determine the stabilized energy eigenfunctions and the first-order corrections to energy for the degenerate energy eigenvectors of H_0 with energy $3\hbar\omega$ when the system experiences the perturbation

$$H_1 = gP_x Q_y .$$

Solution Since H_0 separates, all energy eigenvectors $|n_x, n_y\rangle$ of H_0 and their corresponding energies E^0_n are

$$|n_x, n_y\rangle = |n_x\rangle \otimes |n_y\rangle , \quad E^0_{n_x,n_y} = \hbar\omega\left(n_x + \frac{1}{2}\right) + \hbar\omega\left(n_y + \frac{1}{2}\right) ,$$

where $|n_x\rangle$ and $|n_y\rangle$ are harmonic oscillator energy eigenvectors for the harmonic oscillator in one-dimensional spaces spanned by x and y, respectively, and n_x and n_y independently take the values $0, 1, 2, \ldots$. The three degenerate states with energy $E^0 = 3\hbar\omega$ are, therefore,

$$|1, 1\rangle = |1\rangle \otimes |1\rangle , \quad |0, 2\rangle = |0\rangle \otimes |2\rangle , \quad \text{and} \quad |2, 0\rangle = |2\rangle \otimes |0\rangle .$$

Stabilized solutions $|E^{(0)})$ are of the form

$$|E^{(0)}) = c_1|1,1\rangle + c_2|0,2\rangle + c_3|2,0\rangle\,.$$

The matrix elements of H_1 are readily calculated using (1.2.51) to rewrite P_x in terms of raising operator $a_x^\dagger$ and lowering operator a_x that act in the space $|n_x\rangle$ and writing Q_y in terms of $a_y^\dagger$ and a_y that act in the space $|n_y\rangle$:

$$H_1|n_x, n_y\rangle = ig\sqrt{\frac{m\omega\hbar}{2}}(a_x^\dagger - a_x)|n_x\rangle \otimes \sqrt{\frac{\hbar}{2m\omega}}(a_y^\dagger + a_y)|n_y\rangle\,.$$

With the aid of (1.2.49) and (1.2.50),

$$\begin{aligned}H_1|n_x, n_y\rangle = \frac{i}{2}g\hbar[&\sqrt{n_x+1}\sqrt{n_y+1}|n_x+1, n_y+1\rangle + \sqrt{n_x+1}\sqrt{n_y}|n_x+1, n_y-1\rangle \\ &- \sqrt{n_x}\sqrt{n_y+1}|n_x-1, n_y+1\rangle - \sqrt{n_x}\sqrt{n_y}|n_x-1, n_y-1\rangle]\,.\end{aligned}$$

The matrix elements of H_1 are now readily calculated with the result

$$\begin{aligned}\langle H_1\rangle_{11} &= \langle H_1\rangle_{22} = \langle H_1\rangle_{33} = 0\,,\\ \langle H_1\rangle_{12} &= \langle H_1\rangle_{31} = \frac{i}{\sqrt{2}}g\hbar\,,\\ \langle H_1\rangle_{21} &= \langle H_1\rangle_{13} = -\frac{i}{\sqrt{2}}g\hbar\,,\\ \langle H_1\rangle_{23} &= \langle H_1\rangle_{32} = 0\,.\end{aligned}$$

For this specific problem, the matrix equation (4.3.26) becomes

$$\begin{pmatrix} E^{(1)} - E^0 & -\frac{i}{\sqrt{2}}g\hbar & \frac{i}{\sqrt{2}}g\hbar \\ \frac{i}{\sqrt{2}}g\hbar & E^{(1)} - E^0 & 0 \\ -\frac{i}{\sqrt{2}}g\hbar & 0 & E^{(1)} - E^0 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0\,.$$

Equating the determinant of the above matrix to zero, the following equation is obtained for the energy eigenvalues:

$$\left[E^{(1)} - E^0\right]\left[((E^{(1)} - E^0)^2 - (g\hbar)^2\right] = 0$$

The three normalized stabilized eigenvectors are obtained by substituting each of the three solutions for energy to the above equation into the matrix equation one-at-a-time. The results are tabulated below (Problem 4.18):

Energy	Stabilized Energy Eigenvector
$3\hbar\omega + g\hbar$	$\frac{1}{\sqrt{2}}\left[|1,1\rangle - \frac{i}{\sqrt{2}}|0,2\rangle + \frac{i}{\sqrt{2}}|2,0\rangle\right]$
$3\hbar\omega$	$\frac{1}{\sqrt{2}}\left[|0,2\rangle + |2,0\rangle\right]$
$3\hbar\omega - g\hbar$	$\frac{1}{\sqrt{2}}\left[|1,1\rangle + \frac{i}{\sqrt{2}}|0,2\rangle - \frac{i}{\sqrt{2}}|2,0\rangle\right]$

Note that all three values of $E^{(1)}$ have different values so the perturbation has completely removed the degeneracy.

As the number of degenerate energy eigenvectors increases, it becomes increasingly time-consuming to evaluate the determinant of the matrix in (4.3.26) and find the stabilized energy eigenvectors. If a set of operators $\{B, C, \dots\}$ can be found such that $\{H_0, B, C, \dots\}$ is a complete set of commuting operators and $\{H, B, C, \dots\}$ is also a complete set of commuting operators, then the situation has been reduced to that of the previous section, and the matrix calculation has been avoided entirely. However, even if it is possible to find only a single operator B that commutes with both H_0 and H, where H_0 and B now do not constitute a complete set of commuting operators, the matrix calculation simplifies significantly. Since both the eigenvectors of H and H_0 can be chosen as eigenvectors of B, the stabilized eigenvectors of H with an eigenvalue b_i of B can only be a linear combination of degenerate eigenvectors of H_0 with the same eigenvalue b_i. As will be seen in the following subsection, the matrix equation then reduces to several equations involving smaller matrices that are much easier to solve.

4.3.1 The Stark Effect

The effect of a uniform, static electric field $\mathbf{E}$ on an atomic state was first noted by Stark (1913) and provides an excellent example for the usefulness of perturbation theory. Here the effect of the electric field is determined on the ground state of the hydrogen atom ($n = 1$), which is non-degenerate, and the first excited state of the hydrogen atom ($n = 2$), which is four-fold degenerate.

The Hamiltonian of the hydrogen atom is

$$H_0 = \frac{\mathbf{p}^2}{2\mu} - \frac{e^2}{4\pi\epsilon_0}\frac{1}{Q}, \tag{4.3.27}$$

and describes the relative motion of an electron and proton interacting via the Coulomb force. In the above formula μ is the reduced mass of the electron and proton.

To obtain the operator for the interaction of the electron and proton with the electric field **E**, first note that the dipole moment **d** of the hydrogen atom is

$$\mathbf{d} = -e\mathbf{r}\,, \tag{4.3.28}$$

where e is the charge on the proton and **r** is the position of the electron relative to the proton. (The dipole moment **d** is defined to point from the negative to the positive charge while **r** points from the proton to the electron, accounting for the minus sign in the above equation.) The classical energy of an electric dipole in an electromagnetic field **E** is given by

$$U_{\text{classical}} = -\mathbf{d}\cdot\mathbf{E}\,. \tag{4.3.29}$$

Combining the previous two equations and taking the electric field to point along the z axis,

$$U_{\text{classical}} = -(-e\mathbf{r})\cdot\mathbf{E} = e|\mathbf{E}|r\cos\theta = e|\mathbf{E}|z\,. \tag{4.3.30}$$

From the correspondence principle, the quantum mechanical interaction Hamiltonian is

$$H_1 = U_{\text{quantum}} = e|\mathbf{E}|Q_3\,, \tag{4.3.31}$$

where Q_3 is the 3-component of the relative position operator of the electron with respect to the proton. Including the effect of the electric field, the Hamiltonian H is

$$H_1 = H_0 + e|\mathbf{E}|Q_3\,. \tag{4.3.32}$$

Neglecting spin, the eigenvalues of H_0 are the energy levels of the hydrogen atom as given in (3.2.56). (There the energy eigenvalues are denoted by E_n instead of by E_n^0.) Each energy eigenfunction is a product of a radial wave function and a spherical harmonic. The first few radial wave functions and spherical harmonics are tabulated, respectively, in Tables 3.1 and 3.2.

The energy shift of the ground state resulting from the perturbation (4.3.31) is first calculated. If perturbation theory is to be applicable, (4.2.28) must be satisfied, which implies that the magnitude of the electric field is small ($|\mathbf{E}| \leq 10^3 N/C$). Since the ground state $|n=1,\ell=0,m=0\rangle$ of the hydrogen atom is nondegenerate, the energy to first order in the perturbation can be calculated using (4.2.19):

$$E^{(1)} = E^0 + \langle n=1,\ell=0,m=0|\,e\,|\mathbf{E}|Q_3|n=1,\ell=0,m=0\rangle\,. \tag{4.3.33}$$

Converting the above matrix element into an integral,

$$
\begin{aligned}
E^{(1)} = E^0 + e|\mathbf{E}| \int_{-\infty}^{\infty} \mathrm{d}^3x\, \psi^*_{1,0,0}(\mathbf{x}) z \psi^*_{1,0,0}(\mathbf{x}) \\
= E^0 + e|\mathbf{E}| \int_{-\infty}^{\infty} \mathrm{d}^3x [\psi_{1,0,0}(\mathbf{x})]^2 z\,. \qquad (4.3.34)
\end{aligned}
$$

Making the substitution $z \to -z$, the integral is found to equal the negative of itself and is therefore zero. Consequently, there is no change in the energy level to first-order in the perturbation.

To second-order in the perturbation, the energy $E^{(2)}$ is given by (4.2.23),

$$
E^{(2)}_{n=1} = E^0_{n=1} + \sum_{n',\ell',m' \neq 1,0,0} \frac{|\langle n=1, \ell=0, m=0 | e\, |\mathbf{E}| Q_3 | n', \ell', m' \rangle|^2}{E^0_{n=1} - E^0_{n'}}\,. \qquad (4.3.35)
$$

To evaluate the matrix element, it is converted into an integral,

$$
\langle 1,0,0 | e|\mathbf{E}| Q_3 | n', \ell', m' \rangle = e|\mathbf{E}| \int_0^{\infty} \mathrm{d}r r^2 \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \sin\theta\, \psi_{1,0,0}(\mathbf{x}) z \psi_{n',\ell',m'}(\mathbf{x})\,. \qquad (4.3.36)
$$

Rewriting the hydrogen atom wave functions in terms of the radial wave functions $R_{n,\ell}(r)$ and spherical harmonics $Y^m_\ell(\theta,\phi)$ as given, respectively in Tables 3.1 and 3.2,

$$
\begin{aligned}
&\langle 1,0,0 | e|\mathbf{E}| Q_3 | n', \ell', m' \rangle \\
&= e|\mathbf{E}| \int_0^{\infty} \mathrm{d}r r^3 R^*_{1,0}(r) R_{n',\ell'}(r) \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \sin\theta \left[\cos\theta\, Y^{*0}_0(\theta,\phi) Y^{m'}_{\ell'}(\theta,\phi)\right], \qquad (4.3.37)
\end{aligned}
$$

where $z = r\cos\theta$ has been used. From Table 3.2,

$$
\cos\theta\, Y^{*0}_0(\theta,\phi) = \cos\theta \frac{1}{2\sqrt{\pi}} = \frac{1}{\sqrt{3}} \left[\frac{\sqrt{3}}{2\sqrt{\pi}} \cos\theta\right] = \frac{1}{\sqrt{3}} Y^{*0}_1(\theta,\phi)\,. \qquad (4.3.38)
$$

Substituting the above relation into (4.3.37) and then using the orthonormality relation

$$
\int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \sin\theta\, Y^{*m}_\ell(\theta,\phi) Y^{m'}_{\ell}(\theta,\phi) = \delta_{\ell,\ell'} \delta_{m,m'}\,, \qquad (4.3.39)
$$

It follows that

$$
\begin{aligned}
&\langle 1,0,0|e|\mathbf{E}|Q_3|n',\ell',m'\rangle \\
&= e|\mathbf{E}| \int_0^\infty \mathrm{d}r r^3 R^*_{1,0}(r) R_{n',\ell'}(r) \int_0^{2\pi} \mathrm{d}\phi \int_0^\pi \mathrm{d}\theta \sin\theta \frac{1}{\sqrt{3}} Y_1^{*0}(\theta,\phi) Y_{\ell'}^{m'}(\theta,\phi) \\
&\qquad\qquad = \frac{e|\mathbf{E}|}{\sqrt{3}} \int_0^\infty \mathrm{d}r r^3 R^*_{1,0}(r) R_{n',\ell'}(r) \delta_{\ell',\ell}\delta_{m',0} \, .
\end{aligned}
\tag{4.3.40}
$$

With the aid of the above equation, (4.3.35) becomes

$$
E^{(2)}_{n=1} = E^0_{n=1} + \sum_{n'=2}^{\infty} \frac{e^2|\mathbf{E}|^2}{3} \frac{\left[\int_0^\infty \mathrm{d}r r^3 R^*_{1,0}(r) R_{n',1}(r)\right]^2}{E^0_{n=1} - E^0_{n'}} \, .
\tag{4.3.41}
$$

The integral in the above equation decreases rapidly as the values of n increase with the result that the infinite series converges after summing a relatively small number of terms. Evaluating the first two terms using the explicit expressions for $R_{n,\ell}(r)$ in Table 3.1,

$$
\begin{aligned}
E^{(2)}_{n=1} &= E^0_{n=1} - |\mathbf{E}|^2 4\pi\epsilon_0 r_B^3 \left[\frac{2^{18}}{3^{11}} + \frac{3^8}{2^{15}} + \ldots\right] \\
&= E^0_{n=1} - |\mathbf{E}|^2 4\pi\epsilon_0 r_B^3 \left[1.48 + .20 + \ldots\right] .
\end{aligned}
\tag{4.3.42}
$$

An exact calculation[3] yields the result

$$
E_{n=1} = E^0_{n=1} - \frac{9}{4} |\mathbf{E}|^2 4\pi\epsilon_0 r_B^3 \, .
\tag{4.3.43}
$$

In Problem 4.20 an approximate method is discussed for summing the series that yields a coefficient -2.39 in (4.3.43) instead of the exact result $-9/4 = -2.25$.

As has just been shown, in the presence of an uniform, static electric field, the lowest-order, non-zero energy shift is proportional to the square of the electric field so the effect is called the quadratic Stark effect.

Attention is now directed to the first excited energy level of the hydrogen atom, which is four-fold degenerate. The four zeroth-order eigenvectors are

$$
\begin{aligned}
&|n=2,\ell=1,m=1\rangle, \quad |n=2,\ell=1,m=-1\rangle, \\
&|n=2,\ell=1,m=0\rangle, \quad |n=2,\ell=0,m=0\rangle.
\end{aligned}
\tag{4.3.44}
$$

---

[3] Epstein, Phys. Rev. **28**, 695 (1926); Wentzel, Zeits. fur Phys. **38**, 527 (1926); Waller, *ibid*. **38**, 640 (1926); Van Vleck, Proc. Nat. Acad. Sci. **12**, 662 (1926).

Since the energy levels are degenerate, to simplify the calculation one or more operators are sought that commute with both H_0 and H_1. Because the external field points in the z- or 3-direction, the "exact" Hamiltonian remains invariant under rotations about the z axis but does not remain invariant under rotations about an arbitrary axis. Therefore, H commutes with L_3 but not with $\mathbf{L}^2$,

$$[L_3, H] = 0\,, \qquad \left[\mathbf{L}^2, H\right] \neq 0\,. \tag{4.3.45}$$

Equation (4.3.45) can be checked explicitly in the position representation where the above two equations become, respectively,

$$\left[-i\hbar\frac{\partial}{\partial\phi}\,,\; e|\mathbf{E}|r\cos\theta\right] = 0\,, \tag{4.3.46a}$$

$$\left[-\hbar^2\left\{\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2}\right\},\; e|\mathbf{E}|r\;\cos\theta\right] \neq 0\,. \tag{4.3.46b}$$

Because H commutes with L_3, it is possible to choose eigenstates of H that are also eigenstates of L_3. Since only one of the zeroth-order eigenvectors has an eigenvalue $m = 1$ of L_3 it must be a stabilized eigenvector. Similarly, the single eigenvector with $m = -1$ is also a stabilized eigenvector. Once the zeroth-order approximation for an eigenvector of H is known, the formalism of Sect. 4.2 can be used to calculate the energy shifts. From (4.2.20)

$$E_{n=2,m=\pm1} = E^0_{n=2} + \langle n = 2, \ell = 1, m = \pm1|e|\mathbf{E}|Q_3|n = 2, \ell = 1, m = \pm1\rangle\,. \tag{4.3.47}$$

Converting the above matrix element into an integral,

$$E_{n=2,m=\pm1} = E^0_{n=2} + |e|\mathbf{E}| \int_0^\infty \mathrm{d}^3x\,|\psi_{2,1,\pm1}(\mathbf{x})|^2 z\,. \tag{4.3.48}$$

By making the substitution $z \to -z$, the above integral is found to equal the negative of itself and is therefore zero. To first-order, the degeneracy is not removed for states with $m = \pm1$.

Because there are two eigenvectors of H_0 with $m = 0$, degenerate perturbation theory must be used to determine the stabilized $m = 0$ eigenvectors. The calculation is begun by writing a stabilized eigenvector with $m = 0$ in the form,

$$|E^0_{n=2,m=0}\rangle = c_1|n = 2, \ell = 1, m = 0\rangle + c_2|n = 2, \ell = 0, m = 0\rangle\,. \tag{4.3.49}$$

The diagonal matrix elements of H_1 are equal to zero,

$$\langle H_1\rangle_{11} = \langle n=2, \ell=1, m=0|e|\mathbf{E}|Q_3|n=2, \ell=1, m=0\rangle = 0\,, \tag{4.3.50a}$$

$$\langle H_1\rangle_{22} = \langle n=2, | =0, m=0|e|\mathbf{E}|Q^3|n=2, | =0, m=0\rangle = 0\,, \tag{4.3.50b}$$

for the same reason the diagonal matrix elements in (4.3.48) are zero. The off-diagonal matrix elements are calculated as follows:

$$\langle H_1\rangle_{21}^* = \langle H_1\rangle_{12} = \langle n=2, \ell=1, m=0|e|\mathbf{E}|Q_3|n=2, \ell=0, m=0\rangle\,. \tag{4.3.51}$$

Converting the matrix element into an integral,

$$\langle H_1\rangle_{21}^* = \langle H_1\rangle_{12} = \int_{-\infty}^{\infty} \mathrm{d}^3x\,\psi_{2,1,0}^*(\mathbf{x})\,e\,|\mathbf{E}|\,z\,\psi_{2,0,0}(\mathbf{x})\,. \tag{4.3.52}$$

Writing the hydrogen atom wave functions in the factored form $\psi_{n,\ell,m}(\mathbf{x}) \equiv R_{n,\ell}(r) \times Y_\ell^m(\theta,\phi)$,

$$\begin{aligned}\langle H_1\rangle_{21}^* &= \langle H_1\rangle_{12}\\ &= \int_{-\infty}^{\infty} \mathrm{d}r r^2 \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi R_{2,1}^*(r) Y_1^{*0}(\theta,\phi) e|\mathbf{E}| r\cos\theta R_{2,0}(r) Y_0^0(\theta,\phi).\end{aligned} \tag{4.3.53}$$

The explicit expressions for $R_{n,\ell}(r)$ and $Y_\ell^m(\theta,\phi)$ are given, respectively, in Tables 3.1 and 3.2,

$$\begin{aligned}\langle H_1\rangle_{21}^* = \langle H_1\rangle_{12} &= \int_{-\infty}^{\infty} \mathrm{d}r r^2 \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi \frac{1}{4\sqrt{2\pi}} \frac{1}{(r_B)^{3/2)}} \frac{r}{r_B} e^{-r/2r_B}\\ &\times \cos\theta\, e|\mathbf{E}| r\cos\theta \frac{1}{4\sqrt{2\pi}} \frac{1}{(r_B)^{3/2)}} \left(2 - \frac{r}{r_B}\right) e^{-r/2r_B} = 3e|\mathbf{E}|r_B\,.\end{aligned} \tag{4.3.54}$$

For the problem being solved, the matrix equation (4.3.26) becomes

$$\begin{pmatrix} E^{(1)} - E^0 & -3e|\mathbf{E}|r_B \\ -3e|\mathbf{E}|r_B & E^{(1)} - E^0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0\,. \tag{4.3.55}$$

Equating the determinant of the above matrix to zero, the equation for the energy eigenvalues is

$$(E^{(1)} - E^0)^2 - \langle -3e|\mathbf{E}|r_B\rangle^2 = 0, \quad \text{implying} \quad E^{(1)} = E^0 \pm 3e|\mathbf{E}|r_B\,. \tag{4.3.56}$$

Substituting the above expressions for $E^{(1)}$ into (4.3.55), the corresponding stabilized eigenvectors are found. The energies and corresponding stabilized eigenvectors, two of which were obtained without using degenerate perturbation theory, are listed below:

$$\begin{array}{ll}
\text{Energy} & \text{Stabilized Energy Eigenvector} \\
E^{(1)} = E^0 + 3e|\mathbf{E}|r_B & \frac{1}{\sqrt{2}}[|n=2,\ell=1,m=0\rangle - |n=2,\ell=0,m=0\rangle] \\
E^{(1)} = E^0 - 3e|\mathbf{E}|r_B & \frac{1}{\sqrt{2}}[|n=2,\ell=1,m=0\rangle + |n=2,\ell=0,m=0\rangle] \\
E^{(1)} = E^0 & |n=2,\ell=1,m=1\rangle \\
E^{(1)} = E^0 & |n=2,\ell=1,m=-1\rangle
\end{array} \tag{4.3.57}$$

The factor $1/\sqrt{2}$ has been chosen so that the stabilized eigenvectors are normalized, and overall phase factors have been chosen arbitrarily.

The entire problem could have been solved using degenerate perturbation theory by writing a stabilized eigenvector in the form

$$|E_n^0\rangle = c_1|2,1,0\rangle + c_2|2,0,0\rangle + c_3|2,1,1\rangle + c_4|2,1,-1\rangle\,. \tag{4.3.58}$$

The diagonal matrix elements of H_1 vanish as discussed previously. Also, since L_3 commutes with H_1, all the off-diagonal matrix elements vanish unless both eigenvectors have the same value of m as discussed in Example 4.2.1 on page 216. Using (4.3.54), the matrix equation corresponding to (4.3.26) is then

$$\begin{pmatrix} E^{(1)} - E^0 & 3e|\mathbf{E}|r_B & 0 & 0 \\ 3e|\mathbf{E}|r_B & E^{(1)} - E^0 & 0 & 0 \\ 0 & 0 & E^{(1)} - E^0 & 0 \\ 0 & 0 & 0 & E^{(1)} - E^0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0\,. \tag{4.3.59}$$

The energies and corresponding stabilized eigenfunctions can be calculated from the above equation (Problem 4.21).

To summarize, in the presence of an uniform, static electric field, to first order the energy level $E_{n=2}^0$ of the hydrogen atom splits into three energy levels:

$$E_{n=2}^{(1)} \longrightarrow \begin{cases} E_{n=2}^0 + 3er_B|\mathbf{E}|, & m=0 \\ E_{n=2}^0, & m=\pm 1 \\ E_{n=2}^0 - 3er_B|\mathbf{E}|, & m=0 \end{cases} \tag{4.3.60}$$

The two energy levels with $m = 0$ are non-degenerate while the levels with $m = \pm 1$ are degenerate. The energy shifts are linear in the electric field and are, therefore, called the linear Stark effect.

4.4 Summary

The description of any physical system is only approximate, and only in a limited number of cases can the approximate mathematical description be solved exactly. The most frequently used approximation technique is to split the Hamiltonian H, which only provides an approximate mathematical description of the physical system, into two parts,

$$H = H_0 + H_1 .$$

The eigenvalues and eigenvectors of H_0 are assumed to be known, and H_1 is a "small," time-independent or stationary perturbation. Because the effect of H_1 is "small" compared with that of H_0, the eigenvalues E_n^0 and eigenvectors $|E_n^0\rangle$ of H_0 are almost, respectively, the eigenvalues E_n and eigenvectors $|E_n\rangle$ of H. Stationary perturbation theory provides a systematic method for calculating the small differences $E_n - E_n^0$ and $|E_n\rangle - |E_n^0\rangle$.

The eigenvector of H_0 that is the best approximation to each eigenvector of H is known for either of the following two cases:

Case I A hermitian operator B is known that satisfies

$$[B, H_0] = [B, H] = 0 ,$$

where H_0 and B form a complete set of commuting observables as do H and B.

Case II The eigenvalues of H_0 are non-degenerate.

For the above two cases there are two fundamental equations:

$$E_n = E_n^0 + \frac{\langle E_n^0|H_1|E_n\rangle}{\langle E_n^0|E_n\rangle} ,$$

$$|E_n\rangle = |E_n^0\rangle + \Big[\langle E_n^0|E_n\rangle - 1\Big]\,|E_n^0\rangle + \sum_{n'\neq n} |E_{n'}^0\rangle \frac{\langle E_{n'}^0|H_1|E_n\rangle}{E_n - E_{n'}^0} .$$

The two fundamental equations can be systematically solved for the first-order corrections resulting from H_1, the second-order corrections and so forth. Of special note are the formulas for the energy eigenvalues and energy eigenvectors of H that include first-order corrections resulting from H_1. They are, respectively,

$$E_n^{(1)} = E_n^0 + \langle E_n^0 | H_1 | E_n^0 \rangle \,,$$

$$|E_n^{(1)}\rangle = |E_n^0\rangle + \sum_{n' \neq n} |E_{n'}^0\rangle \frac{\langle E_{n'}^0 | H_1 | E_n^0 \rangle}{E_n^0 - E_{n'}^0} \,.$$

Degenerate, stationary perturbation theory must be used for the remaining case:

Case III The energy eigenvalues of the free Hamiltonian H_0 are degenerate and no hermitian operator B is known that commutes with both H_0 and H such that H_0 and B form a complete set of commuting observables.

The primary difference between Cases I and II and Case III is the following: For Cases I and II the eigenvector $|E_{n,b}, b_i\rangle$ of H is almost the same as the eigenvector $|E_n^0, b_i\rangle$ of H_0. However, for Case III, if there are m degenerate eigenvectors $|E_n^0, b_i\rangle, i = 1, 2, \ldots, m$, of H_0, each of the m corresponding eigenvectors of H is almost the same as some *linear combination* of the m degenerate eigenvectors $|E_n^0, b_i\rangle$. Each linear combination of the $|E_n^0, b_i\rangle$ that is almost an eigenvector of H is called a *stabilized energy eigenvector* $|E_{n,d}^0, d\rangle$ and is written in the form

$$|E_{n,d}^0, d\rangle = \sum_{i=1}^{m} c_i |E_n^0, b_i\rangle \,.$$

The constants c_i and the expressions for energy $E_n^{(1)}$ that include the first-order corrections resulting from H_1 are determined by the matrix equation

$$\begin{pmatrix} E_{n,d}^{(1)} - E_n^0 - \langle H_1 \rangle_{11} & -\langle H_1 \rangle_{12} & -\langle H_1 \rangle_{13} & \cdots \\ -\langle H_1 \rangle_{21} & E_{n,d}^{(1)} - E_n^0 - \langle H_1 \rangle_{22} & -\langle H_1 \rangle_{23} & \cdots \\ -\langle H_1 \rangle_{31} & -\langle H_1 \rangle_{32} & E_{n,d}^{(1)} - E_n^0 - \langle H_1 \rangle_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = 0 \,.$$

In the above equation

$$\langle H_1 \rangle_{ij} \equiv \langle E_n^0, b_i | H_1 | E_n^0, b_j \rangle \,.$$

The effect of a uniform, static electric field on an atomic state is known as the Stark effect. The first non-zero shift in the energy level of the ground-state ($n = 1$) of the hydrogen atom is calculated from second-order perturbation theory and is called the quadratic Stark effect because it is proportional to the square of the magnitude of the electric field. The $n = 2$ energy levels are originally four-fold degenerate. Using first-order degenerate perturbation theory, the $m = \pm 1$ energy levels remain unchanged while one of the $m = 0$ energy levels is shifted upward and the other downward from the unperturbed levels. Since the shift in energy levels is proportional to the magnitude of the electric field, the effect is known as the linear Stark effect.

Problems

For Sect. 4.1.1

4.1 The Hamiltonian of a rotator in a constant magnetic field is given by

$$H = \frac{\mathbf{J}^2}{2I} - g\frac{q}{2m}(B_1 J_1 + B_2 J_2)\,.$$

(a) Verify that the operator

$$J_1' = \frac{1}{\sqrt{B_1^2 + B_2^2}}(B_1 J_1 + B_2 J_2)$$

commutes with H.

(b) Verify that $\mathbf{J}^2 = \mathbf{J}'^2$, where

$$J_1' = \frac{1}{\sqrt{B_1^2 + B_2^2}}(B_1 J_1 + B_2 J_2), \quad J_2' = \frac{1}{\sqrt{B_1^2 + B_2^2}}(B_2 J_1 - B_1 J_2), \quad J_3' = J_3\,.$$

(c) Verify that the J_i' satisfy the algebra of angular momentum.

(d) Using a basis that is an eigenstate of both $\mathbf{J}'^2$ and J_1', determine the exact energy levels as a function of the eigenvalues of $\mathbf{J}'^2$ and J_1'.

For Sect. 4.2

4.2 Consider the following Hamiltonian in one-dimensional space:

$$H = \frac{P^2}{2m} + \frac{1}{2}kQ^2 + \frac{1}{2}bQ^2, \;\; b \ll k\,.$$

(a) Using your knowledge of solutions to the harmonic oscillator, write an exact expression for the energy eigenvalues E_n of H. Expand E_n in powers of b/k neglecting terms of order $(b/k)^n,\; n > 3$.

Now let $H_0 = \frac{P^2}{2m} + \frac{1}{2}kQ^2$ and treat $H_1 = \frac{1}{2}bQ^2$ as a perturbation.

(b) Using second-order stationary perturbation theory, calculate $E_n^{(2)}$. Verify that it agrees with the answer in (a) to the required order in b/k.

(c) Using first-order stationary perturbation theory, express $|E_n^{(1)}\rangle$ in terms of $|E_n^0\rangle$. For the cases $n = 1$ and $n = 2$, verify that

$$|E_0^{(1)}\rangle = |E_0^0\rangle - \frac{1}{4\sqrt{2}}\frac{b}{k}|E_2^0\rangle\,,$$

$$|E_1^{(1)}\rangle = |E_1^0\rangle - \frac{1}{4}\sqrt{\frac{3}{2}}\frac{b}{k}|E_3^0\rangle\,.$$

(d) The exact solutions $|E_0\rangle$ and $|E_1\rangle$ of H satisfy

$$|E_1\rangle = a^\dagger|E_0\rangle\ .$$

Verify that the solutions in (c) are consistent with the above equation to order b/k.

4.3 Let the Hamiltonian of a rotator be given by

$$H = \frac{\mathbf{J}^2}{2I} + \alpha J_3 + \beta J_1\ ,$$

where α and β are constants with $\alpha \gg \beta$. Let $|j, m\rangle$ be an eigenvector of $\mathbf{J}^2$ and J_3 with respective eigenvalues $\hbar^2 j(j+1)$ and $\hbar m$.

(a) Treat $H_1 = \beta J_1$ as a perturbation and determine the energy eigenvalues of the unperturbed Hamiltonian,

$$H_0 = \frac{\mathbf{J}^2}{2I} + \alpha J_3\ .$$

(b) Using stationary perturbation theory with $|j, m\rangle$ as the unperturbed basis, to lowest order in β calculate the first non-zero contribution of $H_1 = \beta J_2$ to the energy eigenvalues.

(c) The exact eigenvalues of H can be determined by defining the following operators:

$$J_1' = J_1,\ J_2' = \frac{1}{\sqrt{\alpha^2+\beta^2}}(\beta J_3 - \alpha J_1),\ J_3' = \frac{1}{\sqrt{\alpha^2+\beta^2}}(\alpha J_3 + \beta J_1)\,.$$

Verify that these "rotated" angular momentum operators satisfy the algebra of angular momentum and that, in terms of the "rotated" operators,

$$H = \eta \mathbf{J}'^2 + \lambda J_3'\ .$$

Express the constants η and λ in terms of the constants I, α and β.

(d) Determine the exact eigenvalues of H and verify that they are consistent with the results obtained in (a) and (b).

4.4 Consider a quantum particle with mass m confined in an infinitely-deep, one-dimensional, potential well. The potential $U(x)$ is given by

$$U(x) = \begin{cases} \infty & x < -a \\ 0 & -a \le x \le a \\ \infty & x > a \end{cases}.$$

The energy eigenvalues E_n^0 and corresponding normalized eigenfunctions $\phi_n(x)$ are as follows:

$$E_n^0 = \frac{n^2\pi^2\hbar^2}{8ma^2} \quad \text{and} \quad \phi_n(x) = \frac{1}{\sqrt{a}}\cos\frac{n\pi x}{2a} \qquad n = 1, 3, 5, \ldots$$

$$E_n^0 = \frac{n^2\pi^2\hbar^2}{8ma^2} \quad \text{and} \quad \phi_n(x) = \frac{1}{\sqrt{a}}\sin\frac{n\pi x}{2a} \qquad n = 2, 4, 6, \ldots$$

A perturbation $U'(x)$ is applied to the system,

$$U'(x) = \begin{cases} U_0 > 0 & -a \le x \le 0\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

The perturbing potential $U'(x)$ in the infinitely-deep, rectangular potential well is shown in Fig. 4.1.

(a) For $n = 1, 3, 5, \ldots$ calculate the exact energy levels to first order in the perturbation $U'(x)$.
(b) For $n = 2, 4, 6, \ldots$ calculate the exact energy levels to first order in the perturbation $U'(x)$.
(c) What inequality must U_0 satisfy if the first-order perturbation calculation is to be meaningful?

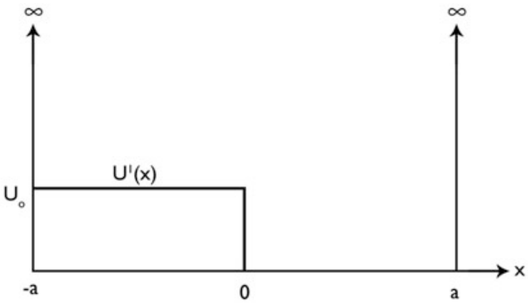


Fig. 4.1 The perturbing potential $U'(x)$ of Problem 4.4 superimposed on the infinitely-deep, rectangular potential well

4.5 Using stationary perturbation theory calculate the lowest-order correction to the ground state energy level of a one-dimensional harmonic oscillator as a result of the perturbation $H_1 = bQ^4$.

4.6 Calculate the lowest order correction to the nth energy level of a one-dimensional harmonic oscillator as a result of the perturbation $H_1 = bQ^2$.

4.7 Calculate the lowest order correction to the n^{th} energy level of a one-dimensional harmonic oscillator as a result of the perturbation $H_1 = bQ^3$.

4.8 Using second-order and first-order stationary perturbation theory, respectively, calculate $E_n^{(2)}$ and $|E_n^{(1)}\rangle$ for a one-dimensional harmonic oscillator subject to a perturbation $H_1 = bP$.

4.9 Calculate the energy of the ground state of the hydrogen atom when it is perturbed by a potential $H_1 = g/(Q_x^2 + Q_y^2 + Q_z^2)$.

4.10 Using first-order perturbation theory calculate the energy eigenvalues and eigenvectors of the Hamiltonian $H = H_0 + H_1$ where

$$H_0 = \frac{\mathbf{J}^2}{2I} - g\frac{q}{2m}B_3 J_3 \quad \text{and} \quad H_1 = \alpha(J_1)^2\,.$$

4.11

(a) For Example 4.2.4 on page 219 calculate $E^{(1)}$ by letting the “exact” Hamiltonian operate on $|E^{(1)}\rangle$ that is given in the text.

(b) Verify to the first order in α that $|E^{(1)}\rangle$ as given in Example 4.2.4 is an eigenstate of the operator

$$B = -g\frac{q}{2m}B_3 J_3 + \alpha(J_2)^2\,.$$

To first order in α, what is the eigenvalue of B when it acts on $|E^{(1)}\rangle$?

4.12 A hydrogen atom experiences a perturbation

$$\langle \mathbf{x}|H_1|\mathbf{x}\rangle = \frac{g}{r^2}\,,\quad g = \text{ constant}\,.$$

(a) Determine two operators that commute with both the free and “exact” Hamiltonians.

(b) For $n = 2$ what are the zeroth-order approximations to the exact energy eigenfunctions?

(c) For $n = 2$ calculate the first-order corrections to energy for each of the energy eigenfunctions found in (b). The following relation is useful:

$$\int_{-\infty}^{\infty} \mathrm{d}^3x\, \psi^*_{n,\ell,m}(\mathbf{x}) \frac{1}{r^2} \psi_{n,\ell,m}(\mathbf{x}) = \frac{1}{(\ell + 1/2)n^3 r_B^2} \,.$$

4.13 Derive the second-order perturbation expression for $E_n^{(2)}$ as given in (4.2.23).

4.14 Derive the second-order perturbation expression for $|E_n^{(2)}\rangle$ as given in (4.2.24).

For Sect. 4.3

4.15 Beginning with (4.3.16) derive the first-order corrections for the energy (4.3.19) and the stabilized energy eigenvectors (4.3.21).

4.16 Verify that the two stabilized energy eigenvectors in (4.3.21) are orthogonal.

4.17 Verify that

$$|s = 1/2, s_1 = \pm 1/2\rangle = \frac{1}{\sqrt{2}} \left[|s = 1/2, s_3 = 1/2\rangle \pm |s = 1/2, s_3 = -1/2\rangle\right] ,$$

where $|s = 1/2, s_1 = \pm 1/2\rangle$ satisfies

$$S_1|s = 1/2, s_1 = \pm 1/2\rangle = \pm\, \frac{\hbar}{2}\, |s = 1/2, s_1 = \pm 1/2\rangle \,.$$

4.18 In Example 4.3.2 on page 232 substitute the explicit expressions for the energy eigenvalues $E^{(1)}$ into the matrix equation for c_1, c_2, and c_3 and determine the three normalized stabilized energy eigenfunctions.

4.19 Using the results from Example 4.3.2, calculate the diagonal matrix elements of H_1 between each of the three stabilized energy eigenvectors. What is the relationship between the three diagonal matrix elements and the three values for $E^{(1)}$? Explain your results.

4.20 Determine an approximate value for the sum in (4.3.35) using the following method: As shown in the text, the only non-zero matrix elements occur for $\ell' = 1$ and $m' = 0$. Writing the first two non-zero terms explicitly,

$$\sum_{n',\ell',m' \neq 1,0,0} \frac{|\langle n = 1, \ell = 0, m = 0|(e|\mathbf{E}|Q_3)|n', \ell', m'\rangle|^2}{E^0_{n=1} - E^0_{n'}}$$

$$= \frac{|\langle 1,0,0|(e|\mathbf{E}|Q_3)|2,1,0\rangle|^2}{E^0_{n=1} - E^0_{n'=2}} + \frac{|\langle 1,0,0|(e|\mathbf{E}|Q_3)|3,1,0\rangle|^2}{E^0_{n=1} - E^0_{n'=3}} + \sum_{n'=4,\ell',m'} \frac{|\langle 1,0,0|(e|\mathbf{E}|Q_3)|n',\ell',m'\rangle|^2}{E^0_{n=1} - E^0_{n'}} \,.$$

From (3.2.56), $E^0_{n'} = E^0_{n=1}/n'^2$. Since the matrix elements rapidly decrease in magnitude for increasing values of n', the value of the above sum is changed only slightly if $E^0_{n'}$ is neglected in the last term. Thus,

$$\sum_{n',\ell',m\neq 1,0,0} \frac{|\langle n=1,\ell=0,m=0|(e|\mathbf{E}|Q_3)|n',\ell',m'\rangle|^2}{E^0_{n=1} - E^0_{n'}}$$

$$\cong \frac{|\langle 1,0,0|(e|\mathbf{E}|Q_3)|2,1,0\rangle|^2}{E^0_{n=1} - E^0_{n=2}} + \frac{|\langle 1,0,0|(e|\mathbf{E}|Q_3)|3,1,0\rangle|^2}{E^0_{n=1} - E^0_{n=3}} + \frac{1}{E^0_{n=1}} \sum_{n'=4,\ell',m'} |\langle 1,0,0|(e|\mathbf{E}|Q_3)|n',\ell',m'\rangle|^2 .$$

(a) Show that

$$\sum_{n'=4,\ell',m'} |\langle 1,0,0|(e|\mathbf{E}|Q_3)|n',\ell',m'\rangle|^2$$

$$= -|\langle 1,0,0|(e|\mathbf{E}|Q_3)|2,1,0\rangle|^2 - |\langle 1,0,0|(e|\mathbf{E}|Q_3)|3,1,0\rangle|^2 + \sum_{n'=0,\ell',m'} |\langle 1,0,0|(e|\mathbf{E}|Q_3)|n',\ell',m'\rangle|^2 \,.$$

(b) Show that

$$\sum_{n'=0,\ell',m'} |\langle 1,0,0|(e|\mathbf{E}|Q_3)|n',\ell',m'\rangle|^2 = |\langle 1,0,0|e|\mathbf{E}|^2(Q_3)^2|1,0,0\rangle \,.$$

Hint: Use the form of (A.4.13) that is appropriate for the hydrogen atom.

(c) Evaluate the following three matrix elements:

$$\langle 1,0,0|e^2|\mathbf{E}|^2(Q_3)^2|1,0,0\rangle$$

$$\langle 1,0,0|e|\mathbf{E}|(Q_3)|2,1,0\rangle$$

$$\langle 1,0,0|e|\mathbf{E}|Q_3|3,1,0\rangle$$

(d) Use your results from (a), (b), and (c) to determine an approximate value for the infinite sum in (4.3.35).

4.21 Beginning with (4.3.59), determine the expressions for $E^{(1)}$ and the corresponding stabilized energy eigenfunctions.

4.22 The Hamiltonian for a harmonic oscillator in two-dimensional space is

$$H_0 = \frac{P_x^2 + P_y^2}{2m} + \frac{1}{2}k(Q_x^2 + Q_y^2) \; .$$

It experiences a perturbation

$$H_1 = gQ_xQ_y, \quad g = \text{ constant.}$$

Determine the stabilized energy eigenvectors and the first-order corrections to the energy for the two degenerate eigenvectors of H_0 with energy $E^0 = 2\hbar\omega$.

4.23 For the Hamiltonian H_0 and the perturbation H_1 of the previous problem, determine the stabilized energy eigenvectors and the first-order corrections to energy for the three degenerate eigenvectors of H_0 with energy $E^0 = 3\hbar\omega$.

Chapter 5
Time Evolution of Quantum Systems

5.1 Time Evolution

5.1.1 Introduction

So far only properties such as the structure and spectra of quantum physical systems have been discussed, properties for which time development is irrelevant and can thus be ignored. In this chapter properties will be discussed that can be understood only by taking into account time development.

Experiments with microphysical systems, which were discussed in detail in Chap. 1, consist of two parts: (1) the preparation procedure that prepares the state using a macroscopic apparatus, and (2) the registration apparatus that registers or detects the observable. Although the experimental apparatus in quantum physics always consists of a preparation apparatus and a registration apparatus, it is not always possible to divide an experiment into preparation and registration parts in a unique way. In the experiment with vibrating CO molecules discussed in Sect. 1.3, the apparatus is represented schematically by Fig. 1.7a and is depicted pictorially in Fig. 1.7b. A monochromator and a collision chamber with a beam of CO molecules prepare the state ρ and constitute the preparation apparatus. The state in this specific experiment is the mixed state $\rho = \sum \rho_n |E_n\rangle\langle E_n|$, (1.3.12); however, it could have also been a pure state $|E_n\rangle\langle E_n|$.

The experiment also consists of the electron multiplier, which is an analyzer and a counter (detector). The analyzer and counter constitute the registration apparatus. The counter registers the electrons that have undergone the energy loss $E_e - E_{e'} = E_m - E_0$, and the observable that represents this energy loss has been denoted $\Lambda_{E_m} = \Lambda_m = |E_m\rangle\langle E_m|$. By changing the magnetic field of the analyzer, the detector can be tuned so that a specific energy loss can be detected, making it possible to detect any one of the observables $\Lambda_0, \Lambda_1, \ldots,$ or $\Lambda_m = |E_m\rangle\langle E_m|$.

By setting the energy of the analyzer to $E_m - E_0$, the observable Λ_{E_m} in the state ρ of the CO molecules prepared by the preparation apparatus of this experiment is registered. As stated in Fundamental Postulate II of quantum theory, the quantities

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9_5

measured in quantum physics are the Born probabilities of an observable Λ. In this case $\Lambda = \Lambda_m = |E_m\rangle\langle E_m|$ in a state ρ. These Born probabilities—also called expectation values of Λ_m in ρ—are defined in the theory by

$$\rho_m^{\text{theoretical}} = \mathrm{Tr}(\rho\Lambda_m) = \mathscr{P}_\rho(\Lambda_m)\,, \tag{5.1.1}$$

and they are measured as ratios of large numbers,

$$\rho_m^{\text{experimental}} = \frac{\text{Number of detector counts with energy loss E}_m}{\text{Number of all detector counts}} = \frac{N_m}{N}\,. \tag{5.1.2}$$

The state could be prepared in a pure state $\rho = |E_n\rangle\langle E_n|$ where $H\,|E_n\rangle = E_n\,|E_n\rangle$. If, for example, $|E_n\rangle = |E_1\rangle$, then $\rho_m = \mathrm{Tr}(|E_1\rangle\langle E_1|\Lambda_m\rangle = \delta_{1,m}$. In general the state operator depends on time, $\rho = \rho(t)$. For the special case of a time-dependent pure state, $\rho(t) = |E_n\rangle\langle E_n|$, the state vector $|E_n\rangle = |E_n(t)\rangle$ depends on time.

For the specific case of the energy loss experiment, the numbers N_m/N measured by the detector refer to the state $\rho(t = 0)$, where $t = 0$ is the time at which the state is prepared. N_m is the number of electrons that have collided with a CO molecule in its ground state with energy E_0 and bumped up the CO molecule into its m-th excited state with energy E_m. This, of course, happens “all the time” as long as the monochromator sends a beam of electrons into the collision chamber where it then collides with the CO molecules. Here “all the time” refers to the time of the clocks in the laboratory. But each time $t_1, t_2, \ldots$ when an individual CO molecule is bumped up into the m-th excited state $|E_m\rangle$ is the initial time $t = 0$ for the specific excited state $|E_m\rangle$ that was created at time t_i. From this time $t = 0$, which is different for each excited state, the excited CO molecule evolves exponentially in time as specified by the decaying state $|E_m(t)\rangle,\ t \geq 0$. But the time-dependence of the state $|E_m(t)\rangle$ is irrelevant for the energy-loss experiment: what the detector measures is the number of electrons N_m per time interval Δt. The number of electrons per unit time $N_m/\Delta t$ that have lost energy $E_m - E_0$ equals the number of the CO molecules per unit time $N_m/\Delta t$ that have been elevated to the m-th energy level.

5.2 Time Evolution of States, Observables and Probabilities

To discuss the question of how Born probabilities change in time, Fundamental Postulates I and II of quantum theory are first reviewed. States and observables are distinguished as follows: States are described by vectors ϕ for pure states and by statistical or density operators ρ (1.3.12) for mixtures. States ϕ and ρ are prepared by a preparation apparatus such as an accelerator.

Observables are described by hermitian operators $A = A^\dagger$. An important, special class of observables are the projection operators $\Lambda = \Lambda^2$. If $\Lambda = |\psi\rangle\langle\psi|$ is a projector onto the one-dimensional space spanned by ψ, then the vector ψ specifies the observable Λ. Observables A are registered by a registration apparatus (detector).

The state vectors ϕ are elements of a scalar-product space Φ and the observables A form an algebra of linear operators in this space. Although often called a Hilbert space, Φ is typically treated as a pre-Hilbert space.[1] Each specific physical system, such as the hydrogen atom or the energy-loss experiment involving CO molecules discussed in Chap. 1, Sect. 1.3, is associated with a specific linear space Φ.

The physical quantities in quantum physics are the probabilities. When a physical system is in the pure state ϕ, the probability of measuring the observable A in this state is denoted $\mathscr{P}_\phi(A)$ and is calculated theoretically as a Born probability,

$$\mathscr{P}_\phi(A) = |\langle A|\phi\rangle|^2 = |\langle \psi(t=0) \rangle A_n |\phi\rangle|^2 \,. \tag{5.2.1}$$

In the above equation the subscript n refers to the nth detector that registers the observable A.

Similarly, when the state is a mixture specified by the statistical or density operator ρ, the Born probability of measuring the observable A is

$$\mathscr{P}_\rho(A) = \mathrm{Tr}(A\rho) = \mathrm{Tr}(A_n\rho) \,. \tag{5.2.2}$$

Experimentally, either when the physical system is a pure state specified by ϕ or a mixture specified by ρ, the probabilities are measured as large count ratios of a detector, which are the relative frequencies,

$$\mathscr{P}_\phi(A) = \frac{N_n}{N} \,. \tag{5.2.3}$$

As in other areas of physics, time development in quantum mechanics can be described by differential equations. However, in quantum mechanics it is possible to represent time development in various equivalent ways or in equivalent "pictures." The Schrödinger and Heisenberg pictures are two of the more commonly used and represent the extremes. In the Schrödinger picture, the state is considered to be time dependent, $\phi = \phi(t)$ and the observable A or Λ is time independent. In the Heisenberg picture, the observable is considered to be time dependent, $A = A(t)$ and $\Lambda = \Lambda(t)$, while the state ϕ is time independent, implying that it remains unchanged for all time.[2]

The comparison of the calculated probabilities and the observed counting rates,

$$\mathscr{P}_{\rho(t)}(A) = \mathscr{P}_\rho(A(t)) \doteqdot \frac{N_n(t)}{N} \,, \tag{5.2.4}$$

---

[1] A pre-Hilbert space is a linear, scalar-product space in which the convergence of infinite sequences $\phi_1,\ \phi_2,\ \ldots \phi_n,\ \phi_{n+1}\ \ldots$ is not discussed. Often the Schwartz space, a space in which convergence is defined by an infinite number of norms, is used for Φ.

[2] The interaction or Dirac picture is a third commonly used picture in which the Hamiltonian H is split into two parts, $H = H_0 + H_1$. The observable A evolves in time with H_0, $A^D(t) = e^{iH_0t/\hbar}A^D(0)e^{-iH_0t/\hbar}$, and the state ϕ evolves in time with H_1, $\phi^D(t) = e^{-iH_1t/\hbar}\phi^D(0)e^{iH_1t/\hbar}$. As must be the case, predictions are the same in each picture: $\mathscr{P}_{\rho^D(t)}(A^D(t)) = \mathscr{P}_{\rho(t)}(A) = \mathscr{P}_\rho(A(t))$.

Table 5.1 Relationship among states, observables and probabilities

State ρ (in-state ϕ^+ of scattering experiment)	Is prepared by	A preparation apparatus (e.g. an accelerator)
Observable A (ψ^- out-observables or "out-states")	Is registered by	A registration apparatus (e.g. a detector)
Experimental quantities are the probabilities of measuring the observable A in the state ρ	These probabilities are calculated theoretically as Born probabilities	Probabilities are measured as ratios of large counts of a detector ("relative frequencies")

tests the theoretical prediction $\mathscr{P}_{\rho(t)}(A) = \mathscr{P}_\rho(A(t))$ and the experimental observation $N_n(t)/N$ (Table 5.1).[3]

The theoretical and experimental expressions for probabilities are as follows:

$$\mathscr{P}_\rho(\Lambda(t)) \equiv \underbrace{\mathrm{Tr}(\Lambda(t)\rho(0)) = \mathrm{Tr}\Lambda(0)\rho(t))}_{\text{theoretical}} \doteqdot \underbrace{N_n/N}_{\text{experimental}} \tag{5.2.5}$$

In the special case $\Lambda = |\psi^-\rangle\langle\psi^-|$ and $\rho = |\phi^+\rangle\langle\phi^+|$,

$$\mathscr{P}_{\phi^+}(\psi^-(t)) \equiv \underbrace{|\langle\psi^-(0)|\phi^+(t)\rangle|^2 = |\langle\psi^-(t)|\phi^+(0)\rangle|^2}_{\text{theoretical}} \doteqdot \underbrace{N_n/N}_{\text{experimental}} \tag{5.2.6}$$

The consequences of Fundamental Postulates I and II are as follows: In quantum physics states are described by density operators, usually denoted by ρ, or, in the special case $\rho = |\phi\rangle\langle\phi|$, by a state vector ϕ. Observables are described by self-adjoint operators $A = A^\dagger$, a special case of which are projection operators $\Lambda = \Lambda^\dagger = \Lambda^2$, or by observable vectors ψ if $\Lambda = |\psi\rangle\langle\psi|$.

Equation (5.2.6) is a special case of the fundamental postulate (5.2.5) for which the observable Λ is given, except for a phase factor, by the vector ψ^-, and the state ρ is given by $\rho = |\phi^+\rangle\langle\phi^+|$. The two extreme cases for describing time evolution are given in Eqs. (5.2.5) and (5.2.6). In the Schrödinger picture, which can be conjectured from classical wave theory, the time dependence of probabilities is described by the time evolution of the state vector $\phi^+(t)$ of the state operator $\rho(t)$ with the observable Λ remaining unchanged in time. All the time dependence resides in the state vector $\phi^+(t)$ so the probability is given by $|\langle\psi^-(0)|\phi^+(t)\rangle|^2$. The Heisenberg picture, which can be conjectured from classical particle mechanics, describes the time dependence of the probabilities as the evolution of the observable

---

[3] The experimental quantity $N_n(t)/N$ necessarily changes in discrete steps while the theoretical quantity $\mathscr{P}_{\rho(t)}(A) = \mathscr{P}_\rho(A(t))$ is a continuous function of t. The comparison between experiment and theory is in principle approximate, but as the number of events N becomes larger, the comparison between experiment and theory becomes more accurate.

$\Lambda(t)$ with the state ρ fixed in time. All the time dependence resides in the observable Λ so the probability is given by $|\langle\psi^-(t)|\phi^+(0)\rangle|^2$.

5.3 Time Evolution in Standard Quantum Mechanics

The mathematical axiom in standard quantum mechanics about the space of states Φ_{state} and the space of observables $\Phi_{\text{observable}}$ is

$$\Phi_{\text{state}} = \Phi_{\text{observable}} \in \mathscr{H} = \text{Hilbert space}\,. \tag{5.3.1}$$

If the Hilbert space axiom (5.3.1) is chosen, then the brackets in (5.2.6) are the scalar products of two vectors $\psi, \phi \in \mathscr{H}$.

Using (5.2.6), the equality $\mathscr{P}_{\psi(t)}(\phi) = \mathscr{P}_{\psi}(\phi(t))$ implies

$$|\langle\phi|\psi(t)\rangle| = |\langle\phi(t)|\psi\rangle|\,. \tag{5.3.2}$$

The equality (5.3.2) can be achieved by choosing a unitary operator[4] $U(t)$, $-\infty < t < +\infty$, to describe time translation of the state,

$$\phi \to \phi(t) = U(t)\phi(0) = U(t)\phi\,, \tag{5.3.3}$$

relative to the observable $\psi = \psi(0)$. In the above equation ϕ represents the preparation apparatus (state), and ψ represents the registration apparatus (detector or observable). Equivalently, one could choose the one-parameter group $U^\dagger(t) = U^{-1}(t)$ of time translation of the observable

$$\psi \to \psi(t) = U^\dagger(t)\,\psi(0) = U^\dagger(t)\,\psi\,, \tag{5.3.4}$$

relative to the state $\phi = \phi(0)$. With this choice the equality (5.3.2) is obvious because

$$\langle\psi\,,\,\phi(t)\rangle = \langle\psi\,,\,U(t)\phi\rangle = \langle U^\dagger(t)\psi\,,\,\phi\rangle = \langle\psi(t)\,,\,\phi\rangle\,, \tag{5.3.5}$$

---

[4] $U(t)$ possesses the following properties:

$$\begin{aligned} &U(0) = I\,,\\ &U^{-1}(t) = U(-t) = U^\dagger(t)\\ &U(t_1+t_2) = U(t_1)Ut_2) = U(t_2)U(t_1)\,;\quad -\infty < t_1\,,t_2 < +\infty\,. \end{aligned}$$

$U(t)$ is a continuous operator function of the parameter t and is called a one-parameter group of unitary operators.

from which (5.3.2) and $\mathscr{P}_{\phi(t)}(\psi) = \mathscr{P}_{\phi}(\psi(t))$ follow.

The unitary operator $U(t) = e^{\mathscr{A}t}$ can be differentiated to yield an operator $\mathscr{A}$ that is a "generator of the unitary transformation":

$$\mathscr{A} = \left.\frac{\mathrm{d}U(t)}{\mathrm{d}t}\right|_{t=0} . \tag{5.3.6}$$

The operator H defined by

$$H \equiv -\frac{\hbar}{i}\mathscr{A} = -\frac{\hbar}{i}\left.\frac{\mathrm{d}U}{\mathrm{d}t}\right|_{t=0} = H^{\dagger} \tag{5.3.7}$$

is self-adjoint. (See Problem 5.1.) When written in terms of H, $U(t)$ is given by

$$U(t) = e^{-iH\,t/\hbar} . \tag{5.3.8}$$

The factor $\hbar$ is Planck's constant divided by 2π and has been included to give H the dimensions of energy. H is then *the time-independent Hamiltonian* of the quantum physical system with states that undergo the time evolution $U(t)$. Each quantum system has an intrinsic Hamiltonian that characterizes the quantum system. For example, the Hamiltonian H of the vibrating CO molecule in its center of mass is given by (1.2.10), and the Hamiltonian of a rotator is given by (2.2.30).

If the Hamiltonian is independent of time, the system is said to be conservative. Using the explicit expression (5.3.8) for $U(t)$ and differentiating the state $\phi(t) = U(t)\phi(0)$, which appears in (5.3.3), with respect to time,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = H\phi(t) , \tag{5.3.9}$$

where $\phi(0)$ is the state at time $t = 0$. Similarly, differentiating the observable $\psi(t) = U^{\dagger}(t)\psi(0)$, which appears in (5.3.4), with respect to time,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi(t) = -H\psi(t) , \tag{5.3.10}$$

where $\psi(0)$ is the observable at time $t = 0$. The usual situation is that the Hamiltonian H is known in terms of the algebra of observables of the quantum mechanical system. Then one wants to determine the time evolution of "all" states $\phi(t)$ or "all" observables $\psi(t)$ that fulfill the dynamical equations (5.3.9) or (5.3.10), respectively. These dynamical equations are introduced as a basic postulate.

5.3.1 Time Development in the Schrödinger Picture—Fundamental Postulate Va

A (conservative) physical system has a generator H of time translation that is a hermitian element of the algebra of observables and is characteristic of the physical system. The time evolution of the state of the physical system is given by (5.3.3) where ϕ is the state of the system at $t = 0$ and $U(t)$ is given by (5.3.8).

The Schrödinger Picture
For a state vector $\phi(t)$, the dynamic equation is the Schrödinger equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = H\phi(t)\,. \tag{5.3.11}$$

To solve this equation one must provide an initial condition such as the state $\phi(0)$ at $t = 0$. In the Schrödinger picture the density operator satisfies the von Neumann equation

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = [H\,,\ \rho(t)]\,, \tag{5.3.12}$$

that governs the time evolution of an initial state $\rho(0)$ at time $t = 0$. (See Problem 5.2.)

However, neither the state vector, the density operator, nor their time evolutions are observable quantities. The observable quantity is the quantum mechanical probability for measuring the observable Λ in the state $\rho(t)$:

$$\mathscr{P}(t) = \mathrm{Tr}(\Lambda\rho(t))\,, \tag{5.3.13}$$

where the observable Λ is kept fixed. For the special case of a pure state $\rho(t) = |\phi(t)\rangle\langle\phi(t)|$ and an observable given by $\Lambda = |\psi\rangle\langle\psi|$:

$$\mathscr{P}(t) = |\langle\psi|\phi(t)\rangle|^2 = \langle\phi|U^\dagger(t)|\psi\rangle\langle\psi|U(t)|\phi\rangle\,. \tag{5.3.14}$$

The probabilities (5.3.13) and (5.3.14) have been written in the Schrödinger picture; they describe the probability of measuring an external, unchanging observable Λ in the time-evolving state $\rho(t)$.

But the final equality in (5.3.14) suggests a second method for describing time evolution. The state ϕ can be taken to be fixed in time while the observable $|\psi\rangle\langle\psi|$ evolves in time according to $U^\dagger(t)|\psi\rangle\langle\psi|U(t)$, yielding a second method, the Heisenberg picture, for describing time development.

5.3.2 Time Development in the Heisenberg Picture—Fundamental Postulate Vb

A (conservative) physical system has a generator H of time translation that is an hermitian element of the algebra of observables and is characteristic of the physical system. The time development of every observable $\Lambda(t)$ of the physical system is given by

$$\Lambda(t) = U^{\dagger}(t)\, \Lambda\, U(t)\,, \tag{5.3.15}$$

where $U(t)$ is given by (5.3.8), and Λ denotes the observable $\Lambda(t)$ at an initial time $t = 0$.

The Heisenberg Picture

The state ρ is considered fixed for all time and the observable $\Lambda(t)$ changes in time according to (5.3.15). The time development of an observable $\Lambda(t)$ is often written in a second form, which is obtained by differentiating (5.3.15):

$$\begin{aligned}\frac{\mathrm{d}\Lambda(t)}{\mathrm{d}t} &= \frac{i}{\hbar}\, He^{itH/\hbar}\Lambda e^{-itH/\hbar} + e^{itH/\hbar}\Lambda\left(-\frac{i}{\hbar}\, H\right)e^{-itH/\hbar} \\ &= \frac{i}{\hbar}(H\,\Lambda(t) - \Lambda(t)\,H) \;= -\frac{i}{\hbar}\,[\Lambda(t)\,,\; H]\,.\end{aligned} \tag{5.3.16}$$

The above equation, called the Heisenberg equation of motion, is solved subject to the initial condition that when $t = 0$, the observable $\Lambda(t)$ is given by the operator

$$\Lambda(t = 0) \equiv \Lambda\,. \tag{5.3.17}$$

Then the probability for the observable $\Lambda(t)$ in the state ρ is given by

$$\mathscr{P}(t) = Tr(\Lambda(t)\rho)\,, \tag{5.3.18}$$

where the state ρ is fixed for all times t.

For the special case that the observable is the "property" $\Lambda_{\psi} \equiv |\psi\rangle\langle\psi|$, which is the projection operator on the one-dimensional subspace spanned by the vector ψ, the Heisenberg differential equation of motion is

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,\psi(t) = \,-H\psi(t)\,. \tag{5.3.19}$$

The equation must be solved subject to the initial condition that the observable at $t = 0$ is in a state given by $\Lambda = |\psi(t = 0)\rangle\langle\psi(t = 0)|$. For this special case of a fixed pure state ϕ and time-evolving observable $|\psi(t)\rangle\langle\psi(t)|$, the probability of detecting the observable $|\psi(t)\rangle\langle\psi(t)|$ in the state ϕ at time t is

$$\mathscr{P}(t) = |\langle\psi(t)|\phi\rangle|^2\,. \tag{5.3.20}$$

5.4 Time-Dependent Hamiltonian

Up to this point the time evolution of quantum systems has been discussed only for systems described by time-independent Hamiltonians. The Hamiltonian does not depend explicitly on time for isolated systems that are also called "conservative" systems. The dynamics of such systems is determined solely by the Hamiltonian, which characterizes the system, is self-adjoint, and is an element of the algebra of observables.

For many interesting and important quantum systems, however, external, time-dependent forces act on the system. For example, when a rotator with a magnetic moment $\boldsymbol{\mu} = g\frac{q}{2m}\mathbf{J}$ is placed in a time-dependent magnetic field $B(t)$, the Hamiltonian $H(t)$ is

$$H(t) = \frac{\mathbf{J}^2}{2I} - \boldsymbol{\mu} \cdot \mathbf{B}(t) = \frac{\mathbf{J}^2}{2I} - g\frac{q}{2m}\mathbf{J} \cdot \mathbf{B}(t)\,. \tag{5.4.1}$$

In Sect. 5.6, a quantum system that has the above form will be discussed in the context of magnetic resonance. For time-dependent Hamiltonians, the time translation operators $U(t)$ for states and $U^\dagger(t)$ for observables are no longer the simple exponentials.

For systems with time-dependent Hamiltonians, time evolution is described by the Schrödinger and Heisenberg equations. In the Schrödinger picture the time evolution of the state vector $\phi(t)$ (or the statistical operator $\rho(t)$) is determined by the Schrödinger equation,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = H(t)\,\phi(t)\,. \tag{5.4.2}$$

Completely equivalently, in the Heisenberg picture the time evolution of the observable $A(t)$ is determined by the Heisenberg equation of motion,

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -\frac{i}{\hbar}[A(t), H(t)] + \frac{\partial A(t)}{\partial t}\,. \tag{5.4.3}$$

The task is now to solve the dynamical equations (5.4.2) or (5.4.3) and then use the solutions to calculate the Born probabilities. For the case where the Hamiltonian is time dependent, a general solution is not known. Furthermore, the procedure for finding the solution depends very much on the specific form of the operator function $H(t)$. In Sect. 5.5 a system is considered for which the Hamiltonian H depends on an external field and is an explicit function of time.

From (5.3.3) it follows that $U(t)$ translates the state in time from time zero to time t, implying that $U(t)$ is the time-translation operator. The operator $-iH/\hbar$, which is proportional to the energy operator H, is the generator of infinitesimal time translations. Similarly, the operators $-iP_i/\hbar$ and $-iJ_i/\hbar$, respectively, generate infinitesimal spatial translations and infinitesimal rotations. Since the Hamiltonian H is hermitian,

$$U(t)^\dagger = (e^{-itH/\hbar})^\dagger = e^{itH/\hbar} = U^{-1}(t)\,. \tag{5.4.4}$$

A linear operator whose adjoint is its inverse is a unitary operator so $U(t)$ is a unitary operator. Time development in quantum mechanics is often described by a unitary operator, and such time development has the following important property: after transforming the state $\phi_S(0)$ into the state $\phi_S(t) = U(t)\phi_S(0)$, it is possible to transform back to the original state $\phi_S(0)$ by applying the operator $U(-t) = U^\dagger(t) = U^{-1}(t)$ (the subscript S stands for Schrödinger). That is,

$$U(-t)\phi_S(t) = e^{itH/\hbar}\phi_S(t) = e^{itH/\hbar}e^{-itH/\hbar}\phi_S(0) = \phi_S(0)\,. \tag{5.4.5}$$

Since it is possible to obtain the initial state at time zero from the state at time t, unitary time development describes processes that are reversible. Later this property will have to be revised.

Example 5.4.1 A harmonic oscillator is in an energy eigenstate $|E_n\rangle$ at time zero. What is its state at time t?

Solution Since the harmonic oscillator is in a state $|E_n\rangle$ at time zero, $\phi_S(0) = |E_n\rangle$. Using (5.3.3) and (5.3.8),

$$|E_n(t)\rangle = U(t)\phi_S(0) = U(t)|E_n\rangle = e^{-itH/\hbar}|E_n\rangle = e^{-itE_n/\hbar}|E_n\rangle\,.$$

The second formulation of time development is expressed as Heisenberg's equation of motion (5.4.3):

$$\frac{\mathrm{d}A_H(t)}{\mathrm{d}t} = -\frac{i}{\hbar}[A_H(t), H] + \frac{\partial A_H(t)}{\partial t} \tag{5.4.6}$$

The Heisenberg equation of motion describes how the observable $A_H(t)$ changes in time just as the Schrödinger equation describes how the state $\phi_S(t)$ changes in time. The subscript H, which stands for Heisenberg, is included to emphasize that the observable $A_H(t)$ is time-dependent. If the physical system is isolated and does not depend on external fields, the observable $A_H(t)$ does not depend explicitly on time, $\frac{\partial A_H}{\partial t} = 0$, and the Heisenberg equation of motion becomes

$$\frac{\mathrm{d}A_H(t)}{\mathrm{d}t} = -\frac{i}{\hbar}[A_H(t), H]\,, \tag{5.4.7}$$

where

$$\frac{\partial A_H(t)}{\partial t} = \frac{\partial H}{\partial t} = 0\,. \tag{5.4.8}$$

The time development of such a physical system, which was discussed in Sect. 3, is determined solely by the intrinsic properties of the system and is not influenced by external factors.

When the Hamiltonian H is not an explicit function of time, $\frac{\partial H}{\partial t} = 0$, the Heisenberg equation of motion can be integrated. This, of course, is the same situation for which the Schrödinger equation can be integrated.

The Schrödinger equation (5.4.2) and the Heisenberg equation of motion (5.4.3) must lead to the same observable facts: a state developing in time according to (5.4.2) must be equivalent to an operator (observable) developing in time according to (5.4.3). The value associated with an experimental measurement cannot depend on the description of time development.

The experimentally measured quantities associated with a state and an observable are, according to Fundamental Postulate II (1.4.3) or, more specifically (1.4.6), the expectation value of the observable A in the state ϕ. That is,

$$\langle A \rangle = \langle \phi | A | \phi \rangle \,. \tag{5.4.9}$$

These numbers will, in general, change in time $\langle A \rangle = \langle A \rangle_t$. One way to describe this change in time is the following: Using the solution $\phi_{\mathrm{S}}(t)$ to the Schrödinger equation (5.4.2) and postulating that the observable A does not change in time,

$$\langle A \rangle_t = \langle \phi_{\mathrm{S}}(t) | A_{\mathrm{S}} | \phi_{\mathrm{S}}(t) \rangle \,. \tag{5.4.10}$$

This is called the Schrödinger picture so the operator A is denoted A_{S}. If (5.3.3) can be used, the above equation can be rewritten

$$\langle A \rangle_t = (\phi_{\mathrm{S}}(0) | U^{\dagger}(t) A_{\mathrm{S}} U(t) | \phi_{\mathrm{S}}(0)) \,. \tag{5.4.11}$$

Another way of describing the change of the expectation value $\langle A \rangle$ in time is the following: Using (5.4.9) and postulating that the state ϕ, called ϕ_{H}, does not change in time, the expectation value of the observable A is

$$\langle A \rangle_t = \langle \phi_{\mathrm{H}} | A_{\mathrm{H}}(t) | \phi_{\mathrm{H}}) \,. \tag{5.4.12}$$

This is called the Heisenberg picture. If the solution (5.3.3) can be used, the above equation can be rewritten

$$\langle A \rangle_t = \langle \phi_{\mathrm{H}} | U^{\dagger}(t) A_{\mathrm{H}}(0) U(t) | \phi_{\mathrm{H}}) \,. \tag{5.4.13}$$

Since the experimental predictions must be the same in either picture, $\langle \phi | A | \phi \rangle$ must be picture-independent implying (5.4.11) and (5.4.13) must be the same. Therefore,

$$\phi_{\mathrm{H}} = \phi_{\mathrm{S}}(0) = U^{\dagger}(t) \phi_{\mathrm{S}}(t) \,, \tag{5.4.14}$$

and

$$A_{\mathrm{S}} = A_{\mathrm{H}}(0) = U(t)A_{\mathrm{H}}(t)U^{\dagger}(t)\,. \tag{5.4.15}$$

Inverting (5.4.15)

$$A_{\mathrm{H}}(t) = U^{\dagger}(t)A_{\mathrm{S}}U(t) = e^{itH/\hbar}A_{\mathrm{S}}e^{-itH/\hbar}\,. \tag{5.4.16}$$

5.4.1 Time Development in the Heisenberg Picture—Fundamental Postulate VIb

A (conservative) physical system has a generator H of time translation that is a hermitian element of the algebra of observables and is characteristic of the physical system. The time evolution of an operator of the physical system is given by (5.4.16) where $A_{\mathrm{S}} = A(0)$ is the operator at $t = 0$ and $U(t)$ is given by (5.3.8).

From (5.4.14) it follows that the state ϕ_{H} in the Heisenberg picture equals $\phi_{\mathrm{S}}(0)$, the state in the Schrödinger picture at time zero. Thus in the Heisenberg picture, the states do not change in time, and all of the time development resides in the operators $A_{\mathrm{H}}(t)$. From (5.4.15), in the Schrödinger picture the operators do not change in time, and all of the time development is contained in the states $\phi_{\mathrm{S}}(t)$.

Assigning all of the time-dependence to the states $\phi_{\mathrm{S}}(t)$ or to the operators $A_{\mathrm{H}}(t)$ is simply convenient mathematically. In the Schrödinger picture the quantum physical system is imagined to be in a different state $\phi_{\mathrm{S}}(t)$ at different times. But the observable, described by the operator A_{S}, remain unchanged for all time. The observable is defined by the apparatus which measures it and by the procedure for the measurement. This apparatus and the procedure for measurement are imagined not to change in time. In the Heisenberg picture, the state of the physical system described by ϕ_{H} is imagined not to change in time. That is, ϕ_{H} describes the entire history of the state of the physical system. The observable A_{H} changes in time so the procedure for measurement changes in time.

To further clarify the difference between the Schrödinger and Heisenberg pictures, a beam pulse passing through a detector is considered. The observable is the property that the pulse is in a particular region of the detector. From the point of view of the Schrödinger picture, the state $\phi_{\mathrm{S}}(t)$ changes as the beam pulse approaches, passes through, and departs the beam detector. The observable A_{S} represents the detector and measurement procedure which do not change in time. The expectation value $\langle A_{\mathrm{S}} \rangle$ is the probability of finding the beam pulse in a particular region of the detector. This probability changes with time. At a time before the beam pulse reaches the detector, the probability is essentially zero. At a time when the "state of the beam" is inside the detector, it is close to unity. After the beam departs the detector, the probability goes to zero again.

How is this same experiment with the beam pulse viewed in the Heisenberg picture? Now the state of the beam pulse remains constant in time, but the probability that the beam is in a certain region changes as the region of detection in the apparatus moves toward the beam pulse. As the detector moves toward the beam pulse, the probability of detecting the beam pulse is essentially zero before the detector reaches it. As the detector continues to move in the same direction, the probability is close to unity when the detector has reached the "state of the beam," which is then inside the detector. Finally the probability drops toward zero as the detector passes through the beam pulse.

In the discussion of the Schrödinger and Heisenberg equations, no distinction was made between the Hamiltonian in each picture. This is because they are the same. Letting $H_{\mathrm{S}} = H$ be the Hamiltonian in the Schrödinger picture, from (5.4.16) it follows that

$$H_{\mathrm{H}}(t) = U^{\dagger}(t) H_{\mathrm{S}} U(t) = e^{itH/\hbar} H e^{-itH/\hbar} = H \,. \tag{5.4.17}$$

To determine when $A_{\mathrm{H}}(t)$ is a solution to the Heisenberg equation of motion (5.4.6), (5.4.16) is differentiated with respect to time,

$$\begin{aligned}\frac{\mathrm{d}A_{\mathrm{H}}(t)}{\mathrm{d}t} &= \left(\frac{i}{\hbar}H + \frac{it}{\hbar}\frac{\partial H}{\partial t}\right) e^{itH/\hbar} A_{\mathrm{H}}(0) e^{-itH/\hbar} \\ &\qquad + e^{itH/\hbar} A_{\mathrm{H}}(0) e^{-itH/\hbar} \left(-\frac{i}{\hbar}H - \frac{it}{\hbar}\frac{\partial H}{\partial t}\right) \\ &\qquad\qquad = -\frac{i}{\hbar}\left[A_{\mathrm{H}}(t), H\right] + \frac{it}{\hbar}\left[\frac{\partial H}{\partial t}, A_{\mathrm{H}}(t)\right] .\end{aligned} \tag{5.4.18}$$

Note that the derivation of the above equation is correct only if $\frac{\partial H}{\partial t}$ commutes with H. Letting $A_{\mathrm{H}} = H$ in (5.4.18),

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = -\frac{i}{\hbar}[H, H] + \frac{it}{\hbar}\left[\frac{\partial H}{\partial t}, H\right] = \frac{it}{\hbar}\left[\frac{\partial H}{\partial t}, H\right] = 0 \,. \tag{5.4.19}$$

Equations (5.4.6) and (5.4.18) are identical only if

$$\frac{\partial A_{\mathrm{H}}(t)}{\partial t} = \frac{it}{\hbar}\left[\frac{\partial H}{\partial t}, A_{\mathrm{H}}(t)\right] , \tag{5.4.20}$$

which, in general, is satisfied only when (5.4.8) is satisfied.

For nonconservative physical systems, in particular systems with time-dependent external forces, the Heisenberg equation of motion for the observable $A_{\mathrm{H}}(t)$ is given by (5.4.6). Taking $A_{\mathrm{H}} = H$ in (5.4.6),

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{i}{\hbar}[H, H] + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \neq 0 \,. \tag{5.4.21}$$

In this case, even if $A_{\mathrm{H}}(t)$ itself does not depend explicitly on time ($\frac{\partial A_{\mathrm{H}}}{\partial t} = 0$), the Hamiltonian still depends explicitly on time ($\frac{\partial H}{\partial t} \neq 0$), and the Heisenberg equation of motion cannot be integrated in the simple way described above. A general theory for integrating (5.4.6) does not exist; however, some systems with explicit time-dependence, such as the one discussed in Sect. 5.5, can be integrated.

One might think that eigenstates of operators in the Schrödinger and Heisenberg pictures develop in time just as the state vectors develop in each picture: this is incorrect. In the Schrödinger picture the eigenvector of A_{S}, which has an eigenvalue a_n at time $t = 0$, is denoted $|a_n, 0\rangle$

$$A_{\mathrm{S}}|a_n, 0\rangle = a_n|a_n, 0\rangle \,. \tag{5.4.22}$$

Since A_{S} does not change in time, the above relation is true for all time. Thus in the Schrödinger picture the eigenvectors do not change in time although the states do.

In the Heisenberg picture the eigenvector of $A_{\mathrm{H}}(t = 0) \equiv A_{\mathrm{H}}(0) = A_{\mathrm{S}}$ with eigenvalue a_n is denoted $|a_n, 0\rangle$ and satisfies

$$A_{\mathrm{H}}(0)|a_n, 0\rangle = a_n|a_n, 0\rangle \,. \tag{5.4.23}$$

The vector $|a_n, t\rangle$ defined by

$$|a_n, t\rangle = U^{\dagger}(t)|a_n, 0\rangle = e^{itH/\hbar}|a_n, 0\rangle \,, \tag{5.4.24}$$

is now shown to be an eigenvector of $A_{\mathrm{H}}(t)$ with the same eigenvalue. Using (5.4.16),

$$\begin{aligned} A_{\mathrm{H}}(t)|a_n, t\rangle &= e^{itH/\hbar} A_{\mathrm{H}}(0) e^{-itH/\hbar} e^{itH/\hbar}|a_n, 0\rangle \\ &= e^{itH/\hbar} A_{\mathrm{H}}\langle 0)|a_n, 0\rangle = a_n e^{itH/\hbar}|a_n, 0\rangle = a_n|a_n, t\rangle. \end{aligned} \tag{5.4.25}$$

In the Heisenberg picture, the states do not change in time but the eigenvectors of observables do. Comparing (5.4.25) with (5.3.3), the eigenvectors in the Heisenberg picture develop in “the opposite direction” as compared to the states in the Schrödinger picture.

To further understand the time development of eigenvectors in each picture, the expectation value of an operator A in the state ϕ is again examined. Writing ϕ in terms of orthonormal eigenvectors $|a_n\rangle$ of the operator A (See, for example, (A.4.13).),

$$\phi = \sum_n |a_n\rangle\langle a_n|\phi\rangle. \tag{5.4.26}$$

Using this result, the expectation value of A in the state ϕ becomes,

$$\langle A\rangle = \langle\phi|A|\phi\rangle = \sum_m\sum_n \langle\phi|a_m\rangle\langle a_m|A|a_n\rangle\langle a_n|\phi\rangle = \sum_m\sum_n \langle\phi|a_m\rangle a_n\langle a_m|a_n\rangle\langle a_n|\phi\rangle$$
$$= \sum_m\sum_n \langle\phi|a_m\rangle a_n\delta_{m,n}\langle a_n|\phi\rangle = \sum_n a_n\langle\phi|a_n\rangle\langle a_n|\phi\rangle\,. \qquad (5.4.27)$$

Since the expectation value of an operator is the sum of each possible eigenvalue a_n times the probability of obtaining that value,

$$\text{Probability of obtaining the value } a_n = \langle\phi|a_n\rangle\langle a_n|\phi\rangle = |\langle a_n|\phi\rangle|^2. \qquad (5.4.28)$$

In the Schrödinger picture,

$$\text{Probability of obtaining the value } a_n \text{ at time } t = |\langle a_n, 0|\phi_{\rm S}(t)\rangle|^2 \qquad (5.4.29)$$

Using (5.3.3) and (5.3.8),

$$\text{Probability of obtaining the value } a_n \text{ at time } t = |\langle a_n, 0|e^{-itH/\hbar}|\phi_{\rm S}(0)\rangle|^2 \qquad (5.4.30)$$

But from (5.4.14), $\phi_{\rm S}(0) = \phi_{\rm H}$. Therefore,

$$\text{Probability of obtaining the value } a_n \text{ at time } t = |\langle a_n, 0|e^{-itH/\hbar}|\phi_{\rm H}\rangle|^2. \qquad (5.4.31)$$

In the Heisenberg picture,

$$\text{Probability of obtaining the value } a_n \text{ at time } t = |\langle a_n, t|\phi_{\rm H}\rangle|^2. \qquad (5.4.32)$$

Using (5.4.24),

$$\text{Probability of obtaining the value } a_n \text{ at time } t = |\langle a_n, 0|e^{-itH/\hbar}\phi_{\rm H}\rangle|^2. \qquad (5.4.33)$$

The result (5.4.30) from the Schrödinger picture agrees with the result (5.4.33) from the Heisenberg picture as it must.

Two physically equivalent pictures that describe time development have been discussed: the Schrödinger picture and the Heisenberg picture. Depending on the specific problem, one picture may be more convenient than the other for calculations, but the predictions of each are identical. In addition to the two extreme pictures discussed, which place all the time dependence either in the states (Schrödinger picture) or in the operators (Heisenberg picture), there are other possible pictures. In fact, any two pictures related by a unitary transformation are equivalent physically.

Example 5.4.2 Let the states and operators in pictures 1 and 2 be denoted, respectively, by ϕ_1, A_1 and ϕ_2, A_2. If

$$\phi_2 = U^\dagger \phi_1 , \tag{5.4.34a}$$

$$A_2 = U^\dagger A_1 U , \tag{5.4.34b}$$

where U is a unitary operator, show that the expectation value of the operator A is the same in either picture.

Solution The expectation value of A in picture 2, denoted $\langle A \rangle_2$, is given by

$$\langle A \rangle_2 = \langle \phi_2 | A_2 | \phi_2 \rangle .$$

Using (5.4.34a),

$$\langle A \rangle_2 = \langle \phi_1 | U A_2 U^\dagger | \phi_1 \rangle .$$

Using (5.4.34b),

$$\langle A \rangle_2 = \langle \phi_1 | A_1 | \phi_1 \rangle = \langle A \rangle_1 .$$

The transformations relating the Heisenberg picture and the Schrödinger picture are of the form (5.4.34), where $U = e^{-itH/\hbar}$.

Although it will not discussed in detail here, a third picture, usually called the interaction picture, is of great importance. In this picture the Hamiltonian H is split into two parts, $H = H^{(0)} + H'$ where the "free" Hamiltonian is $H^{(0)}$ and the "interaction" Hamiltonian is H'. Denoting the states and operators in the interaction picture, respectively, by ϕ_{ip} and A_{ip}, the interaction picture and the Schrödinger picture are related by the unitary transformations

$$\phi_{\text{ip}}(t) = U^{(0)\dagger}(t)\phi_{\text{S}}(t) , \tag{5.4.35a}$$

$$A_{\text{ip}}(t) = U^{(0)\dagger}(t) A_{\text{S}} U^{(0)}(t) , \tag{5.4.35b}$$

where $U^{(0)}(t) = e^{-itH_{\text{S}}^{(0)}/\hbar}$. The interaction picture is especially useful because the time development of the operators in the interaction picture is determined by the "free" Hamiltonian in the interaction picture. (See Problem 5.8.)

When the Hamiltonian is not an explicit function of time, the relationships among the Schrödinger, Heisenberg, and interaction pictures are summarized in Table 5.2 on the next page.

From Table 5.2 on the facing page note that in the interaction picture the time development of the state is determined by the interaction Hamiltonian in the interaction picture, $H'_{\text{ip}} = U^{(0)^\dagger}(t) H'_{\text{S}} U^{(0)}(t)$. (See Problem 5.9.) Since $U^{(0)}(t) =$

Table 5.2 Relationships among the Schrödinger, Heisenberg and interaction pictures

Picture	State	Operator
Schrödinger	$\phi_S(t) = U(t)\phi_S(0)$	A_S
Heisenberg	$\phi_H = \phi_S(0)$	$A_H(t) = U^\dagger(t)A_S U(t)$
Interaction	$\phi_{ip}(t) = U^{(0)\dagger}(t)\phi_S(t)$	$A_{ip}(t) = U^{(0)\dagger}(t)A_S U^{(0)}(t)$
Picture	Equation of motion for states	
Schrödinger	$i\hbar\frac{d\phi_S(t)}{dt} = H\phi_S(t)$	
Heisenberg	$\phi_H = \phi_S(0) =$constant	
Interaction	$i\hbar\frac{d\phi_{ip}(t)}{dt} = H'_{ip}\phi_{ip}(t)$	
Picture	Equation of motion for operators	
Schrödinger	$\frac{dA_S}{dt} = \frac{\partial A_S}{\partial t}$	
Heisenberg	$\frac{dA_H(t)}{dt} = -\frac{i}{\hbar}\left[A_H(t), H\right] + \frac{\partial A_H(t)}{\partial t}$	
Interaction	$\frac{dA_{ip}(t)}{dt} = -\frac{i}{\hbar}\left[A_{ip}(t), H_{ip}^{(0)}\right] + \frac{\partial A_{ip}(t)}{\partial t}$	

$e^{-itH_S^{(0)}/\hbar}$ does not in general commute with H'_S, it follows that in general $H'_{ip} \neq H'_S$. Because $[H_S^{(0)}, U^{(0)}(t)] = 0$,

$$H_{ip}^{(0)} = U^{(0)\dagger}(t)H_S^{(0)}U^{(0)}(t) = U^{(0)\dagger}(t)U^{(0)}(t)H_S^{(0)} = H_S^{(0)}\,, \tag{5.4.36}$$

so the "free" Hamiltonian in the interaction picture is the same as the free Hamiltonian in the Schrödinger picture.

5.5 Precession of a Spinning Particle in a Magnetic Field: The Interpretation of the Schrödinger and Heisenberg Pictures

This section provides a specific example illustrating the relationship between the Schrödinger and Heisenberg pictures for time evolution of a quantum system. Additionally, it also serves as an introduction to magnetic resonance that is discussed in the next section.

As an example of time development, the motion of a rotator, a quantum mechanical system with rotational but no translational degrees of freedom, is described in a uniform, external magnetic field. The Hamiltonian of this system is given by

$$H = \frac{1}{2I}\mathbf{J}^2 - \boldsymbol{\mu}\cdot\mathbf{B}(t)\,, \tag{5.5.1}$$

where I is the moment of inertia, $\mathbf{J}$ is the angular momentum, and $\boldsymbol{\mu}$ is the magnetic dipole moment of the rotator in a uniform magnetic field $\mathbf{B}$. Although the magnetic field may depend on time t, in this section it is chosen to be time-independent. The problem is first solved in the Schrödinger picture and then in the Heisenberg picture to further clarify the differences and similarities of the two pictures. While the primary purpose is to illustrate the relationship between the two pictures, the motion of spin in a magnetic field is of considerable practical interest in connection with magnetic resonance experiments. Such experiments, which involve time-dependent magnetic fields, will be discussed in the next section.

5.5.1 Precession of a Classical Spinning Particle in a Magnetic Field

Before considering the precession of a spinning quantum system using the Schrödinger and Heisenberg pictures, the corresponding classical system, consisting of a charged mass-point with orbital angular momentum $\mathbf{l}$ in a uniform, constant magnetic field, is solved. Classically, the torque $\boldsymbol{\tau}$ equals the time rate of change of the angular momentum $\mathbf{l}$,

$$\boldsymbol{\tau} = \frac{\mathrm{d}\,\mathbf{l}}{\mathrm{d}\,t}\,. \tag{5.5.2}$$

When the charged mass-point rotates in the magnetic field, a uniform field does not exert a net magnetic force; however, it does exert a net torque $\boldsymbol{\tau}$,

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}, \tag{5.5.3}$$

on the magnetic dipole moment $\boldsymbol{\mu}$. The magnetic moment $\boldsymbol{\mu}$ of a point-charge q rotating with angular momentum $\mathbf{l}$ is given by

$$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{l}. \tag{5.5.4}$$

The expression for the magnetic dipole moment $\boldsymbol{\mu}$ of the object was discussed in Chap. 4, Sect. 4.1 (See (4.1.3).). In general, a particle with spin $\mathbf{l}$ has a magnetic dipole moment

$$\boldsymbol{\mu} = g\frac{e}{2m}\mathbf{l}\,, \tag{5.5.5}$$

where g is the Landé factor for the particle. Combining (5.5.2)–(5.5.5), the following differential equation for orbital angular momentum is obtained:

$$\frac{\mathrm{d}\mathbf{l}}{\mathrm{d}t} = \boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B} = g\frac{e}{2m}\mathbf{l} \times \mathbf{B} \tag{5.5.6a}$$

In component form the above equation becomes

$$\frac{\mathrm{d}l_i}{\mathrm{d}t} = g\frac{e}{2m}\epsilon_{ijk}l_j B_k \,. \tag{5.5.6b}$$

Equation (5.5.6) implies that the change in orbital angular momentum is perpendicular both to the magnetic field and to the orbital angular momentum itself.

For convenience the z axis is chosen to point in the direction of **B**. Then the three components of the vector equation (5.5.6) are as follows:

$$\frac{\mathrm{d}\,l_x}{\mathrm{d}t} = \frac{geB}{2m}\,l_y \tag{5.5.7a}$$

$$\frac{\mathrm{d}\,l_y}{\mathrm{d}t} = -\frac{geB}{2m}\,l_x \tag{5.5.7b}$$

$$\frac{\mathrm{d}\,l_z}{\mathrm{d}t} = 0 \tag{5.5.7c}$$

The solution to (5.5.7c) is immediate: the z-component of orbital angular momentum remains constant in time,

$$l_z(t) = l_z(0) \,. \tag{5.5.8}$$

In particular, if the initial condition

$$l_z(0) = 0 \,, \tag{5.5.9a}$$

is chosen, then

$$l_z(t) = 0 \,, \tag{5.5.9b}$$

for all t.

To solve (5.5.7a) and (5.5.7b), a trial solution is used that has the form

$$l_x = c_1 \cos \omega t + c_2 \sin \omega t \,, \tag{5.5.10a}$$

$$l_y = c_3 \sin \omega t + c_4 \cos \omega t \,, \tag{5.5.10b}$$

Substituting (5.5.10) into (5.5.7a) and (5.5.7b), (5.5.10) is a solution provided

$$c_3 = -c_1, \;\; c_2 = c_4, \tag{5.5.11}$$

and

$$\omega = \omega_L \,, \quad \text{where} \quad \omega_L = \frac{geB}{2m} \,, \tag{5.5.12}$$

is the Larmor frequency of the spinning particle with the Landé factor g. (The frequency $\omega_L = qB/2m$ is called the classical Larmor frequency.)

Initial conditions determine the values of the constants c_1 and c_4. Specifically, from (5.5.10) when $t = 0$, $l_x = c_1$ and $l_y = c_4$ or

$$c_1 = l_x(0), \;\; c_4 = l_y(0). \tag{5.5.13}$$

Using (5.5.10)–(5.5.13), the solution to (5.5.7) is

$$l_x(t) = l_x(0) \, \cos\omega_L t + l_y(0) \, \sin\omega_L t, \tag{5.5.14a}$$

$$l_y(t) = -l_x(0) \, \sin\omega_L t + l_y(0) \, \cos\omega_L t, \tag{5.5.14b}$$

$$l_z(t) = l_z(0). \tag{5.5.14c}$$

From (5.5.14) it follows that

$$\mathbf{l}^2(t) = l_x^2(t) + l_y^2(t) + l_z^2(t) = l_x^2(0) + l_y^2(0) + l_z^2(0) = \mathbf{l}^2(0). \tag{5.5.15}$$

Thus the length of the orbital angular momentum vector remains constant as it moves in the magnetic field. Furthermore, denoting the angle between the magnetic field **B** and the angular momentum $\mathbf{l}(t)$ by α,

$$\cos\alpha = \frac{\mathbf{B} \cdot \mathbf{l}(t)}{|\mathbf{B}| \cdot |\mathbf{l}(t)|} = \frac{B_z l_z(t)}{B_z |\mathbf{l}(t)|} = \frac{l_z(0)}{\sqrt{l_x^2(0) + l_y^2(0) + l_z^2(0)}}, \tag{5.5.16}$$

implying that the angle α between **B** and $\mathbf{l}(t)$ remains constant. Note that when $l_z(0) = 0$, cos $\alpha = 0$ or $\alpha = \pi/2$, implying that the angular momentum **l** is always perpendicular to the z axis and, therefore, rotates in the $x - y$ plane.

To obtain a clear geometrical understanding of the motion of $\mathbf{l}(t)$ in the presence of a constant external magnetic field **B**, the time $t = 0$ is chosen when $l_y = 0$. At this instant

$$l_x(t=0) = l_x(0), \quad l_y(t=0) = 0, \quad l_z(t=0) = l_z(0), \tag{5.5.17}$$

and (5.5.14) becomes

$$l_x(t) = l_x(0) \cos\, \omega_L t, \tag{5.5.18a}$$

$$l_y(t) = -l_x(0) \sin\, \omega_L t, \tag{5.5.18b}$$

$$l_z(t) = l_z(0). \tag{5.5.18c}$$

The situation at $t = 0$ is depicted in Fig. 5.1a on the next page. From (5.5.18), a quarter of a period T later, $\omega_L T/4 = \pi/2$, implying that $\cos(\frac{\omega_L T}{4}) = 0$ and $\sin(\frac{\omega_L T}{4}) = 1$. Thus $l_x(t = \frac{T}{4}) = 0$, $l_y(t = \frac{T}{4})) = -l_x(0)$ and $l_z(t = \frac{T}{4})) = l_z(0)$.

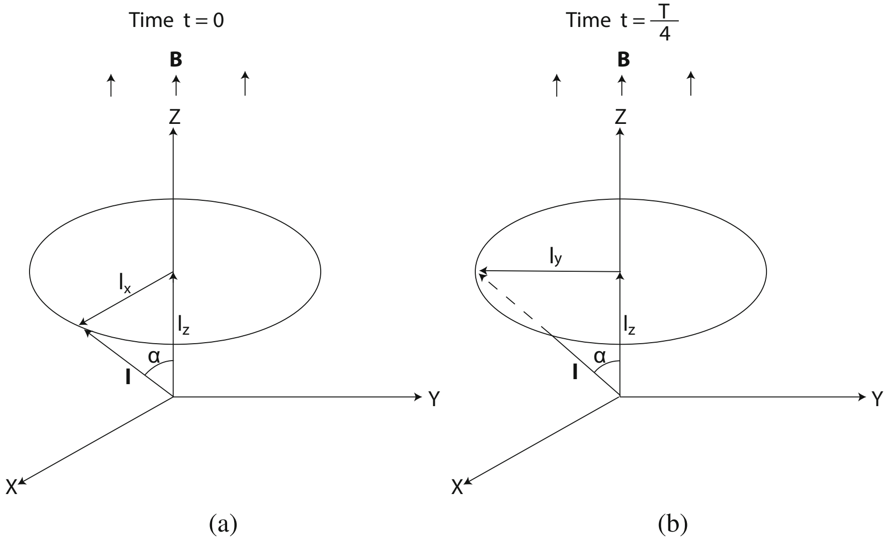


Fig. 5.1 Motion of the angular momentum **l** of a charged mass-point in a constant, uniform magnetic field **B**

This situation is depicted in Fig. 5.1b. From Fig. 5.1 it follows that the tip of the angular momentum vector is rotating in a circle of radius $l_x(0)$ about the z axis. When viewed from above, the direction of rotation is clockwise. Classically, the orbital angular momentum vector precesses about the z axis.

5.5.2 Mathematical Preliminaries: Rotation of Operators and States

In quantum mechanics it is possible to view the time development in different ways. If the Schrödinger picture is used, the state $\phi(t)$ precesses about the z axis and the angular momentum operators are constant in time. If the Heisenberg picture is used, the observables evolve in time with the result that the angular momentum operators $J_i = J_i(t)$ precess about the z axis while the state does not change. Because the quantum solutions involve either rotating vectors or rotating operators, before the solutions can be obtained, it is necessary to further discuss such rotations.

The unit vector $\mathbf{e}_3$, which points in the z direction, has been distinguished physically by placing it in the direction of the magnetic field, $\mathbf{B} = B\,\mathbf{e}_3$, so rotations about the 3 axis are of particular interest. As shown in Fig. 5.2 on the following page, when the unit vector $\mathbf{e}_1$ is rotated by an angle θ about the 3 axis, it becomes the new vector $\mathbf{n}$ where

$$\mathbf{e}_1 \to \mathbf{n} = \mathbf{e}_1 \cos\theta + \mathbf{e}_2 \sin\theta. \tag{5.5.19}$$

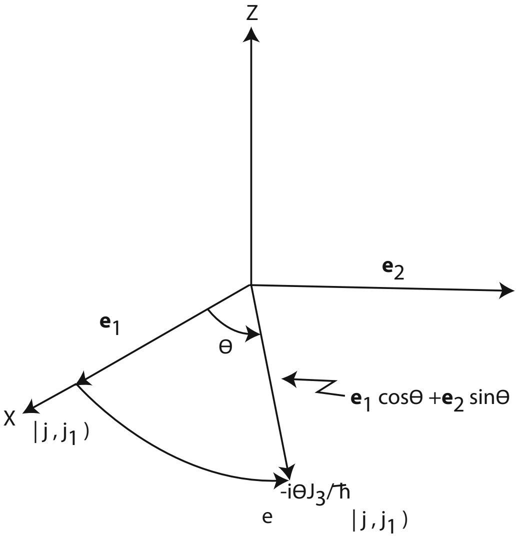


Fig. 5.2 Rotation of the unit vector $\mathbf{e}_1$ about the z axis by an angle θ

For every rotation in three-dimensional space, there is a corresponding operator in the space $\mathscr{H}$ of quantum mechanical states and observables such that the transformation (5.5.19) is represented by a transformation in $\mathscr{H}$. When $\mathbf{e}_1$ is transformed into $\mathbf{n}$, the component of $\mathbf{J}$ along $\mathbf{e}_1$, namely $J_1 = \mathbf{J} \cdot \mathbf{e}_1$, is transformed into the component of $\mathbf{J}$ in the direction of $\mathbf{n}$, which is $\mathbf{J} \cdot \mathbf{n}$,

$$J_1 \to \mathbf{J} \cdot \mathbf{n} = \mathbf{J} \cdot (\mathbf{e}_1 \cos\theta + \mathbf{e}_2 \sin\theta) = J_1 \cos\theta + J_2 \sin\theta\,. \tag{5.5.20}$$

The eigenvectors $|j, j_1\rangle$ that satisfy

$$\begin{aligned} \mathbf{J}^2 |j, j_1\rangle &= \hbar^2 j(j+1)|j, j_1\rangle, \\ J_1 |j, j_1\rangle &= \hbar j_1 |j, j_1\rangle, \end{aligned} \tag{5.5.21}$$

also are rotated. (Usually an eigenvector of $\mathbf{J}^2$ and $J_z \equiv J_3$ is chosen, but this new basis system, which will be employed later, is an eigenvector of J_1 instead of J_3.)

As will be shown, the unitary operator $U_\theta = e^{-i\theta J_i/\hbar}$ rotates an eigenvector by an angle θ about the i axis, the direction being determined by the right-hand rule. Thus under the rotation (5.5.19), the eigenvector $|j, j_1\rangle$ becomes

$$|j, j_1\rangle \to e^{-i\theta J_3/\hbar}|j, j_1\rangle. \tag{5.5.22}$$

Since $e^{-i\theta J_3/\hbar}|j, j_1\rangle$ is the eigenvector obtained by rotating $|j, j_1\rangle$ by an angle θ about the z axis, the vector $e^{-i\theta J_3/\hbar}|j, j_1\rangle$ must be an eigenstate of the angular momentum operator in the direction $\mathbf{n}$, which is $\mathbf{J} \cdot \mathbf{n}$, with eigenvalue $\hbar j_1$. That is, as a result of the rotation, $J_1 \to \mathbf{J} \cdot \mathbf{n}$ where

$$\mathbf{J} \cdot \mathbf{n} e^{-i\theta J_3/\hbar}|j, j_1\rangle = \hbar j_1 e^{-i\theta J_3/\hbar}|j, j_1\rangle\,. \tag{5.5.23}$$

Using (5.5.20),

$$(J_1\cos\theta + J_2\sin\theta)e^{-i\theta J_3/\hbar}|j, j_1\rangle = \hbar j_1 e^{-i\theta J_3/\hbar}|j, j_1\rangle. \tag{5.5.24}$$

Since

$$\begin{aligned}(e^{-i\theta J_3/\hbar}J_1 e^{i\theta J_3/\hbar})e^{-i\theta J_3/\hbar}|j, j_1\rangle &= e^{-i\theta J_3/\hbar}J_1|j, j_1\rangle,\\ &= \hbar j_1 e^{-i\theta J_3/\hbar}|j, j_1\rangle\,,\end{aligned} \tag{5.5.25}$$

the operator obtained by rotating J_1 by an angle θ about the 3 axis according to the right-hand rule is $e^{-i\theta J_3/\hbar}J_1 e^{i\theta J_3/\hbar}$. Comparing (5.5.24) and (5.5.25), under the rotation (5.5.19),

$$J_1 \to e^{-i\theta J_3/\hbar}J_1 e^{i\theta J_3/\hbar} = J_1\cos\theta + J_2\sin\theta = \mathbf{J}\cdot\mathbf{n}. \tag{5.5.26}$$

The above equation can be proved directly by using the mathematical identity

$$e^{iB}Ae^{-iB} = A + i[B, A] + \frac{(i)^2}{2!}[B, [B, A]] + \frac{(i)^3}{3!}[B, [B, [B, A]]] + \cdots\,. \tag{5.5.27}$$

(The proof of the above identity is sketched in Problem 5.6.) Taking $A = J_1$ and $B = -\theta\, J_3/\hbar$, (5.5.27) becomes

$$\begin{aligned}e^{i(-\theta J_3/\hbar)}J_1 e^{-i(-\theta J_3/\hbar)} &= J_1 + i\left[-\frac{\theta}{\hbar}J_3, J_1\right] + \frac{(i)^2}{2!}\left[-\frac{\theta}{\hbar}J_3, \left[-\frac{\theta}{\hbar}J_3, J_1\right]\right]\\ &\quad + \frac{(i)^3}{3!}\left[-\frac{\theta}{\hbar}J_3, \left[-\frac{\theta}{\hbar}J_3, \left[-\frac{\theta}{\hbar}J_3, J_1\right]\right]\right] + \cdots\end{aligned} \tag{5.5.28}$$

With the help of the commutation relations

$$[J_3, J_1] = i\hbar J_2, \quad [J_3, J_2] = -i\hbar J_1,$$

(5.5.28) becomes

$$\begin{aligned}e^{-i\theta J_3/\hbar}J_1 e^{i\theta J_3/\hbar} &= J_1 + \theta J_2 - \frac{\theta^2}{2!}J_1 - \frac{\theta^3}{3!}J_2 + \cdots\\ &= J_1\left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \cdots\right) + J_2\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \cdots\right)\\ &= J_1\cos\theta + J_2\sin\theta.\end{aligned} \tag{5.5.29}$$

To obtain the last line of the above expression, the Taylor series expansions of cos θ and sin θ have been used. Equation (5.5.29) proves that the operator obtained from J_1 by rotating by an angle θ about the 3 axis is given by (5.5.26). Similarly, rotating J_2 by an angle θ about the 3 axis yields (See Problem 5.16.)

$$J_2 \to e^{-i\theta J_3/\hbar} J_2 e^{i\theta J_3/\hbar} = -J_1 \sin\theta + J_2 \cos\theta \,. \tag{5.5.30}$$

Example 5.5.1 The angular momentum operator J_1 is rotated 90° about the 3 axis in a counterclockwise direction when viewed from above.

(a) Using Fig. 5.2 on page 272 express the rotated operator in terms of J_1, J_2 and J_3.
(b) Establish the result in (a) mathematically.

Solution

(a) From Fig. 5.2 when the operator J_1 is rotated 90° it becomes the operator J_2.
(b) This result is established mathematically by using the specific value $\theta = \pi/2$ in (5.5.29):

$$J_1 \to e^{-i(\pi/2)J_3/\hbar} J_1 e^{i(\pi/2)J_3/\hbar} = J_1 \cos(\pi/2) + J_2 \sin(\pi/2) = J_2 \,.$$

The value of an observable in the three-dimensional space in which an experiment takes place is determined by the experimental apparatus. In the mathematical space $\mathscr{H}$, the observables are represented by operators J_i or $|j, j_i\rangle\langle j, j_i|$, where $i = 1, 2, 3$. The expectation value of the observable $|j, j_1\rangle\langle j, j_1|$, for example, is the probability that the value of the angular momentum operator J_1 is j_1:

$$\begin{aligned}\text{Probability of the value } j_1 \text{ in the state } \phi(t) &= \langle\phi(t)|j, j_1\rangle\langle j, j_1|\phi(t)\rangle \\ &= |\langle\phi(t)|j, j_1\rangle|^2 \,. \end{aligned} \tag{5.5.31}$$

5.5.3 *Precession of a Spinning Particle in a Magnetic Field: The Schrödinger Picture*

Having completed the mathematical preliminaries, a quantum mechanical description of the precession of angular momentum in a constant, uniform external magnetic field $\mathbf{B} = B\,\mathbf{e}_3$ will be given in terms of probabilities and expectation values, first using the Schrödinger picture and then using the Heisenberg picture. The Hamiltonian is given by

$$H = \frac{\mathbf{J}^2}{2I} - \boldsymbol{\mu} \cdot \mathbf{B} \,, \tag{5.5.32}$$

where **B** is the constant, uniform external magnetic field and the magnetic dipole operator $\boldsymbol{\mu}$ is

$$\boldsymbol{\mu} = g\frac{e}{2m}\mathbf{J}\,. \tag{5.5.33}$$

In the above equation g is the Landé factor, e is the charge of the rotator, and m is its mass.

The energy levels of the Hamiltonian (5.5.32) were determined in Chap. 4, Sect. 4.1. By choosing the z-direction as the direction of **B**, the z-component of angular momentum J_3 commutes with the Hamiltonian H. Then it is possible to choose as a basis system $|j, j_3\rangle$, which are simultaneous eigenstates of H and J_3. The energy eigenvalues are immediately found to be

$$H|j, j_3\rangle = \left(\frac{\mathbf{J}^2}{2I} - g\frac{e}{2m}J_3 B\right)|j, j_3\rangle = \left(\frac{\hbar^2 j(j+1)}{2I} - g\frac{e\hbar}{2m}j_3\, B\right)|j, j_3\rangle\,. \tag{5.5.34}$$

In the Schrödinger picture the state changes according to

$$\phi(t) = e^{-iHt/\hbar}\,\phi(t=0)\,, \tag{5.5.35}$$

and the angular momenta observables J_i and the observables[5] $|j, j_i\rangle\langle j, j_i|$ remain constant in time.

In the classical calculation earlier in this section, at the particular time $t = 0$ the angular momentum was assumed to point along the x axis, implying that $l_x(0) = l_x$, $l_y(0) = 0$, and $l_z(0) = 0$. In an analogous manner for the quantum calculation, at $t = 0$ the state of the system is assumed to be an eigenstate of J_1 with eigenvalue $\hbar\, j_1$. Thus the state vector at $t = 0$ is assumed to be

$$\phi(t=0) = |j, j_1\rangle. \tag{5.5.36}$$

After choosing the 3 axis to point in the direction of **B**, it is possible to calculate $\phi(t)$ using (5.5.35), (5.5.36) and the explicit expression for H given in (5.5.32),

$$\begin{aligned}\phi(t) = e^{-iHt/\hbar}\,\phi(t=0) &= (e^{-i[\hbar j(j+1)/2I]t}e^{ig(e/2m)B\,t\,J_3/\hbar})|j, j_1\rangle \\ &= e^{-i[\hbar j(j+1)/2I]t}e^{i\omega_L\,t J_3/\hbar}|j, j_1\rangle\,. \end{aligned}\tag{5.5.37}$$

---

[5] The operators $|j, j_i\rangle\langle j, j_i|$ represent the observables that measure the probability for the angular momentum values j, j_i. That is, $|\langle\phi(t)|j, j_i\rangle|^2$ is the probability for the values j, j_i in the state $\phi(t)$.

The frequency ω_L is the Larmor frequency (5.5.12).[6]

The first exponential term on the right-hand side of (5.5.37) is a phase factor that depends on time. Recalling that $e^{-i\theta J_3/\hbar}$ rotates a state about the 3 axis by an angle θ according to the right-hand rule (counterclockwise as viewed from the positive 3 axis), the second exponential factor $e^{i\omega_L tJ_3/\hbar}$ rotates the state by an angle $\omega_L t$ in a clockwise direction as viewed from the positive 3 axis.

The expectation values of the three components of angular momentum are now calculated in the state $\phi(t)$. According to (5.5.36), initially the (pure) state $\phi(0)$ is such that the magnetic moment (of all rotators in the ensemble) has a component in the positive 1-direction equal to $g(e/2m)\, j_1$. The expectation value of the operator J_1 at time $t = 0$ is

$$\langle \phi(t=0)|J_1|\phi(t=0)\rangle = \langle j, j_1|J_1|j, j_1\rangle = \hbar j_1. \tag{5.5.38}$$

Using (5.5.37) and then (5.5.29), the expectation value of J_1 at time t is

$$\begin{aligned}\langle \phi(t)|J_1|\phi(t)\rangle &= \langle e^{-iHt/\hbar}\phi(0)|J_1|e^{-iHt/\hbar}\phi(0)\rangle = \langle j, j_1|e^{-i\omega_L tJ_3/\hbar}J_1 e^{i\omega_L tJ_3/\hbar}|j, j_1\rangle\\ &= \langle j, j_1|J_1 \cos\omega_L t + J_2 \sin\omega_L t|j, j_1\rangle = \hbar j_1 \cos\omega_L t\,. \end{aligned}\tag{5.5.39}$$

The final line of the above equation was obtained using

$$\langle j, j_1|J_1|j, j_1\rangle = \hbar j_1, \quad \langle j, j_1|J_2|j, j_1\rangle = 0\,. \tag{5.5.40}$$

(See Problem 5.13.) Using (5.5.30), a similar calculation yields

$$\langle \phi(t)|J_2|\phi(t)\rangle = -\hbar j_1 \sin\omega_L t. \tag{5.5.41}$$

The relation

$$\langle \phi(t)|J_3|\phi(t)\rangle = 0 \tag{5.5.42}$$

follows immediately from the formula $[J_3\,, e^{i\omega_L tJ_3/\hbar}] = 0$, the initial condition $\phi(0) = |j, j_1\rangle$, and

$$\langle j, j_1|J_3|j, j_1\rangle = 0, \tag{5.5.43}$$

(See Problem 5.13.)

In the classical model the components of orbital angular momentum were determined as a function of time in (5.5.18). Statements about observables in classical mechanics are replaced by statements about expectation values of observables in

---

[6] The Larmor frequency ω_L for a proton (g = 2.79, $e = 1.60 \times 10^{-19}$ C, $m = 1.67 \times 10^{-27}$ kg) in a typical magnetic field $B = 1.0$ T is $\omega_L = 1.34 \times 10^8$ Hz, which is in the radio frequency range.

quantum mechanics. The quantum equations (5.5.39), (5.5.41), and (5.5.42), which specify the expectation values of angular momentum operator as a function of time, are identical in form to the classical equations (5.5.18) that give the classical orbital angular momentum as a function of time.

Because the operators J_i are time-independent in the Schrödinger picture, the eigenvectors of J_i are also constant in time. From (5.5.31), the probability of measuring the value $\hbar j_1'$ for the observable J_1 at time t is equal to the expectation value of the projection operator $|j, j_1'\rangle\langle j, j_1'|$ in the state $\phi(t)$. That is,

$$\begin{aligned}\mathscr{P}_{\phi(t)}(|j, j_1'\rangle\langle j, j_1'|) &= \langle\phi(t)|j, j_1'\rangle\langle j, j_1'|\phi(t)\rangle \\ &= \langle j, j_1|e^{-i\omega_L t J_3/\hbar}|j, j_1'\rangle\langle j, j_1'|e^{i\omega_L t J_3/\hbar}|j, j_1\rangle\,. \end{aligned} \tag{5.5.44}$$

5.5.4 Precession of a Spinning Particle in a Magnetic Field: The Heisenberg Picture

To describe the same system in the Heisenberg picture, the time development of the angular momentum operators must be determined, which necessitates solving the Heisenberg equations of motion (5.4.3),

$$\frac{\mathrm{d}J_i}{\mathrm{d}t} = -\frac{i}{\hbar}[J_i, H]\,. \tag{5.5.45}$$

Taking the 3 axis to be in the direction of the magnetic field and using (5.5.32) and (5.5.33), the above equation becomes

$$\frac{\mathrm{d}J_i}{\mathrm{d}t} = -\frac{i}{\hbar}\left[J_i, \frac{\mathbf{J}^2}{2I} - g\frac{e}{2m}BJ_3\right] = i\,\frac{geB}{2m\hbar}[J_i, J_3]\,. \tag{5.5.46}$$

The second equality above follows because $[J_i, \mathbf{J}^2] = 0$. Using the commutation relations (2.3.1) for angular momentum, the equations (5.5.46) for J_1, J_2 and J_3 become, respectively,

$$\frac{\mathrm{d}J_1}{\mathrm{d}t} = \frac{geB}{2m}J_2\,, \tag{5.5.47a}$$

$$\frac{\mathrm{d}J_2}{\mathrm{d}t} = -\frac{geB}{2m}J_1\,, \tag{5.5.47b}$$

$$\frac{\mathrm{d}J_3}{\mathrm{d}t} = 0. \tag{5.5.47c}$$

The Heisenberg equations (5.5.47) have the same form as the classical equations (5.5.7) and can be solved similarly with the result

$$J_1(t) = J_1(0)\cos\omega_L t + J_2(0)\sin\omega_L t = e^{iHt}J_1(0)e^{-iHt}\,, \tag{5.5.48a}$$

$$J_2(t) = -J_1(0)\sin\omega_L t + J_2(0)\cos\omega_L t = e^{iHt}J_2(0)e^{-iHt}\,, \tag{5.5.48b}$$

$$J_3(t) = J_3(0) = e^{iHt}J_3(0)e^{-iHt}\,. \tag{5.5.48c}$$

In the above equation the Larmor frequency $\omega_L = geB/2m$, and the final equalities in (5.5.48a) and (5.5.48b) follow, respectively, from (5.5.29) and (5.5.30).

In the Heisenberg picture, the state does not change in time. Here, just as in the Schrödinger picture, the state ϕ is assumed to be such that at $t = 0$ the probability of obtaining the value $\hbar j_1$ is unity in a measurement of $J_1(0)$. That is,

$$\mathscr{P}_{\phi(t=0)}(|j, j_1\rangle\langle j, j_1|) = \langle\phi(t=0)|j, j_1\rangle\langle j, j_1|\phi(t=0)\rangle = 1. \tag{5.5.49}$$

Since the system is being described in the Heisenberg picture where the states are time-independent, the above equation is equivalent to saying that, except for a possible phase factor, the state $\phi = \phi(0)$ for all times is given by

$$\phi = |j, j_1\rangle\,. \tag{5.5.50}$$

The expectation values for the components $J_i(0)$ in this state are

$$\langle\phi|J_1(0)|\phi\rangle = \langle j, j_1|J_1(0)|j, j_1\rangle = \hbar j_1\,, \tag{5.5.51a}$$

$$\langle\phi|J_2(0)|\phi\rangle = \langle j, j_1|J_2(0)|j, j_1\rangle = 0\,, \tag{5.5.51b}$$

$$\langle\phi|J_3(0)|\phi\rangle = \langle j, j_1|J_3(0)|j, j_1\rangle = 0\,. \tag{5.5.51c}$$

Equations (5.5.51) are identical to (5.5.40) and (5.5.43). (The derivation is sketched in Problem 5.13.) Using (5.5.48) and (5.5.51), the expectation value of the components of angular momentum at time t are found to be

$$\langle\phi|J_1(t)|\phi\rangle = \langle j, j_1|J_1(0)\cos\omega_L t + J_2(0)\sin\omega_L t|j, j_1\rangle = \hbar j_1\cos\omega_L t\,, \tag{5.5.52a}$$

$$\langle\phi|J_2(t)|\phi\rangle = \langle j, j_1| - J_1(0)\sin\omega_L t + J_2(0)\cos\omega_L t|j, j_1\rangle = -\hbar j_1\sin\omega_L t\,, \tag{5.5.52b}$$

$$\langle\phi|J_3(t)|\phi\rangle = \langle j, j_1|J_3(0)|j, j_1\rangle = 0. \tag{5.5.52c}$$

The expectation values for angular momentum calculated above in the Heisenberg picture agree with the expectation values (5.5.39), (5.5.41), and (5.5.42) calculated in the Schrödinger picture.

In the Heisenberg picture observables are functions of time, implying that the angular momentum operators $J_i(t) = e^{iHt/\hbar} J_i(0) e^{-iHt/\hbar}$ depend on time. Also, from (5.4.24), for systems described by time-independent Hamiltonians, the observable eigenvector $|j, j_1\rangle$ at time t is, except for a possible phase factor, given by

$$|j, j_1, t\rangle = e^{itH/\hbar} |j, j_1\rangle . \tag{5.5.53}$$

For the specific Hamiltonian (5.5.32)

$$\begin{aligned}|j, j_1, t\rangle &= e^{i[\hbar j(j+1)/2I]t} e^{-ig[eB/2m]t J_3/\hbar} |j, j_1\rangle \\ &= e^{i[\hbar j(j+1)/2I]t} e^{-i\omega_L t J_3/\hbar} |j, j_1\rangle .\end{aligned} \tag{5.5.54}$$

For the observable J_1, the probability of detecting the value $\hbar j_1'$ is the expectation value of the projection operator $|j, j_1', t\rangle\langle j, j_1', t|$ in the state ϕ and is given by

$$\mathscr{P}_\phi(|j, j_1', t\rangle\langle j, j_1', t|) = \langle \phi | j, j_1', t\rangle\langle j, j_1', t | \phi\rangle . \tag{5.5.55}$$

Using the fact that $\phi = |j, j_1\rangle$ has been chosen to be an eigenstate of J_1 with and eigenvalue j_1, from (5.5.54) it follows that

$$\mathscr{P}_\phi(|j, j_1', t\rangle\langle j, j_1', t|) = \langle j, j_1 | e^{-i\omega_L t J_3/\hbar} | j, j_1'\rangle\langle j, j_1' | e^{i\omega_L t J_3/\hbar} | j, j_1\rangle . \tag{5.5.56}$$

The probability calculated above in the Heisenberg picture agrees with the result (5.5.44) in the Schrödinger picture as it must.

From the preceding discussion it follows that predictions for probabilities are the same in either the Schrödinger or the Heisenberg picture. (The predictions are probabilities that are measured as detector counts in the observation apparatus that is placed relative to the state, which is the preparation apparatus.) The only difference in the two pictures is the way in which the experiment is visualized.

To further clarify the difference between the Schrödinger and Heisenberg pictures, let ϕ represent the state with all magnetic moments aligned along the $\mathbf{e}_1$-direction with $j_1 = j$. This situation is shown schematically in Fig. 5.3a on the following page by the arrow pointing along $\mathbf{e}_1$. The observable is described by the detector. When the detector is placed along the $\mathbf{e}_1$ axis, it registers magnetic moments pointing in the $\mathbf{e}_1$-direction. The observable is the probability of detecting a magnetic moment directed along the $\mathbf{e}_1$ axis. Mathematically the observable is described by $|j, j_1 = j\rangle\langle j, j_1 = j|$. Figure 5.3a shows the experimental situation at time $t = 0$.

In the Schrödinger picture the state $\phi(0) = |j, j_1\rangle$ "rotates" according to (5.5.37). After a time t the direction of the magnetic moments of the state have rotated about the 3 axis by an angle $-\omega_L t$. Since the observable (operator) $|j, j_1 = j\rangle\langle j, j_1 = j|$ does not change in time, the situation is as shown in Fig. 5.3b for $\omega_L t = \pi/2$.

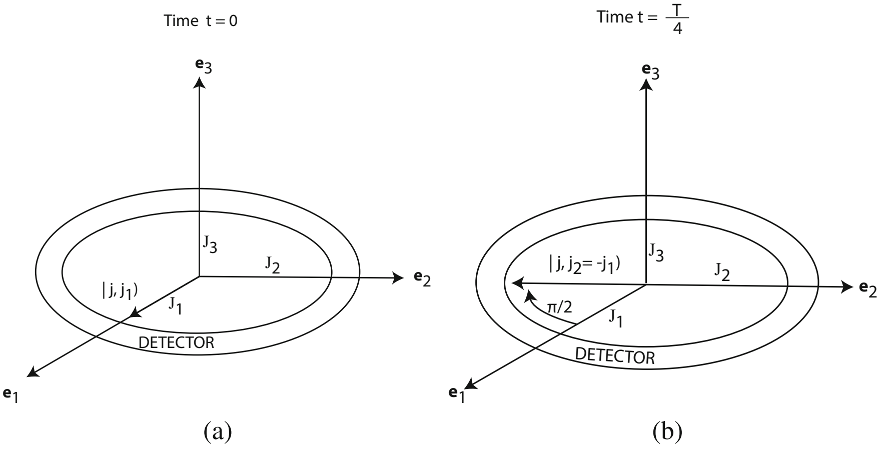


Fig. 5.3 The precession of the angular momentum from the point of view of the Schrödinger picture. During one quarter of a period, the state is rotated $\pi/2$ in a clockwise direction

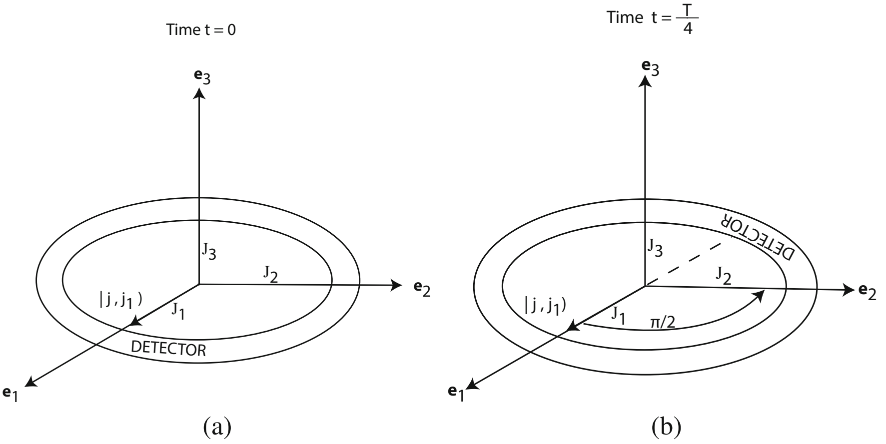


Fig. 5.4 The precession of the angular momentum from the point of view of the Heisenberg picture. During one quarter of a period, the detector is rotated $\pi/2$ in a counterclockwise direction

In the Heisenberg picture the observable (operator) $|j, j_1 = j, t\rangle\langle j, j_1 = j, t|$ changes in time according to (5.5.54). Figure 5.4a on the current page again shows the situation at time $t = 0$. After a time t, the detector has rotated about the 3 axis by an angle $+\omega_L t$. Because the state ϕ does not change in time, the situation is that shown in Fig. 5.4b for $\omega_L t = \pi/2$.

In comparing Figs. 5.3b and 5.4b, the observed probabilities (detector counts) are the same because space is homogeneous for rotations, The result of the calculations (5.5.44) and (5.5.55) confirms that they are indeed identical.

The difference between Figs. 5.3b and 5.4b is strictly in the mind. In the Schrödinger picture (Fig. 5.3b), the experiment is viewed from the perspective of the detector and the preparation apparatus is seen to rotate clockwise when viewed from the positive 3 axis. In the Heisenberg picture (Fig. 5.4b), the experiment is viewed from the perspective of the preparation apparatus and the detector is seen to rotate in a counterclockwise direction when viewed from the positive 3 axis.

Matrix elements of the type that appear in (5.5.49) are especially easy to calculate when the spin $j = 1/2$.

Example 5.5.2 Verify that if, in addition to the usual angular momentum commutation relations (2.3.1), the angular momentum operators also satisfy

$$(J_i J_j + J_j J_i) \equiv \{J_i, J_j\} = \frac{\hbar^2}{2} I \delta_{ij}; \quad i, j = 1, 2, 3 = x, y, z, \tag{5.5.57}$$

then the only allowed value of total angular momentum is $j = 1/2$. (The symbol $\{A, B\}$ is called an anticommutator.)

Solution From (5.5.57) it follows that

$$J_1^2 = J_2^2 = J_3^2 = \frac{\hbar^2}{4} I. \tag{5.5.58}$$

With this result,

$$\hbar^2 j(j+1)|j, j_3\rangle = \mathbf{J}^2|j, j_3\rangle = (J_1^2 + J_2^2 + J_3^2)|j, j_3\rangle = \hbar^2 \frac{3}{4}|j, j_3\rangle\,.$$

Comparing the first and last expressions in the above equation, the only allowed value of total angular momentum is $j = 1/2$.

Example 5.5.3 Show that when spin $j = 1/2$, operators of the form $e^{-i\theta J_3/\hbar}$ that appear in (5.5.56) can be written in the form

$$e^{-i\theta J_3/\hbar} = I \cos\frac{\theta}{2} - 2i\frac{J_3}{\hbar}\sin\frac{\theta}{2}, \quad j = 1/2 \text{ only}. \tag{5.5.59}$$

Solution Expanding $e^{-i\theta J_3/\hbar}$ in a Taylor series,

$$e^{-i\theta J_3/\hbar} = 1 - i\frac{\theta}{\hbar}J_3 + \frac{1}{2!}\left(-i\frac{\theta}{\hbar}J_3\right)^2 + \frac{1}{3!}\left(-i\frac{\theta}{\hbar}J_3\right)^3$$

$$+ \frac{1}{4!}\left(-i\frac{\theta}{\hbar}J_3\right)^4 + \frac{1}{5!}\left(-i\frac{\theta}{\hbar}J_3\right)^5 + \cdots = I\left[1 - \frac{1}{2!}\left(\frac{\theta}{2}\right)^2 + \frac{1}{4!}\left(\frac{\theta}{2}\right)^4 - \cdots\right]$$

$$- 2i\frac{J_3}{\hbar}\left[\frac{\theta}{2} - \frac{1}{3!}\left(\frac{\theta}{2}\right)^3 + \frac{1}{5!}\left(\frac{\theta}{2}\right)^5 - \cdots\right],$$

where the property (5.5.58) has been used. The sum of the terms in the first and second square brackets are, respectively, the Taylor series expansions for $\cos(\theta/2)$ and $\sin(\theta/2)$, yielding the desired result.

Example 5.5.4 Calculate $e^{-i(\theta_1 J_1+\theta_2 J_2+\theta_3 J_3)/\hbar}$ when $j = 1/2$.

Solution Rather than directly expanding the exponent in a Taylor series, the exponent is first rewritten in a more compact form. Defining the unit vector

$$\mathbf{n} = \frac{1}{\theta}(\theta_1\mathbf{e}_1 + \theta_2\mathbf{e}_2 + \theta_3\mathbf{e}_3) \quad \text{where} \quad \theta = \sqrt{\theta_1{}^2 + \theta_2{}^2 + \theta_3{}^2}\,,$$

and

$$\tilde{J}_3 = \mathbf{n}\cdot\mathbf{J} = \frac{1}{\theta}(\theta_1 J_1 + \theta_2 J_2 + \theta_3 J_3)\,,$$

it follows that

$$e^{-i(\theta_1 J_1+\theta_2 J_2+\theta_3 J_3)/\hbar} = e^{-i\theta\tilde{J}_3/\hbar}\,.$$

Using (5.5.59),

$$\begin{aligned} e^{-i(\theta_1 J_1+\theta_2 J_2+\theta_3 J_3)/\hbar} &= I\cos\left(\frac{\theta}{2}\right) - 2i\frac{\tilde{J}_3}{\hbar}\sin\left(\frac{\theta}{2}\right) \\ &= I\cos\left(\frac{\theta}{2}\right) - 2i\,\frac{\mathbf{n}\cdot\mathbf{J}}{\hbar}\sin\left(\frac{\theta}{2}\right), \quad j = 1/2 \text{ only.} \end{aligned} \tag{5.5.60}$$

Restricting to the case where the spin $j = 1/2$, the probability (5.5.44) or, equivalently, (5.5.56) is evaluated using (5.5.60). Specifically, here the probability $\mathscr{P}_\phi(|1/2, j_1' = -1/2\langle\rangle 1/2, j_1' = -1/2|) \equiv \mathscr{P}_{j_1'=-1/2}$ is calculated, which is the probability for measuring the value $j_1' = -1/2$ if the value of $J_1(0)$ at time $t = 0$ is, with certainty, $j_1 = 1/2$:

$$\mathscr{P}_{j_1'=-1/2} = |\langle 1/2, j_1' = -1/2|e^{i\omega_L t J_3/\hbar}|1/2, j_1 = 1/2\rangle|^2\,. \tag{5.5.61}$$

Using (5.5.59),

$$\mathscr{P}_{j_1'=-1/2} = |\langle 1/2, j_1' = -1/2|I\cos\frac{\omega_L t}{2} + 2i\frac{J_3}{\hbar}\sin\frac{\omega_L t}{2}|1/2, j_1 = 1/2\rangle|^2\,. \tag{5.5.62}$$

Now,

$$\langle j = 1/2, j_1' = -1/2 | j = 1/2, j_1 = 1/2\rangle = 0,$$

$$\langle j = 1/2, j_1' = -1/2 | J_3 | j = 1/2, j_1 = 1/2\rangle = \frac{i\hbar}{2},$$

as calculated in Problem 5.14. Thus

$$\mathscr{P}_{j_1'=-1/2} = \sin^2 \frac{\omega_L t}{2}. \tag{5.5.63}$$

Similarly, the probability of measuring a value $\langle J_1\rangle = 1/2$ is

$$\mathscr{P}_{j_1'=1/2} = \cos^2 \frac{\omega_L t}{2}. \tag{5.5.64}$$

The sum of the probabilities for measuring $\langle J_1\rangle = -1/2$ and $\langle J_1\rangle = 1/2$ is unity as it must be. Since the expectation value of an operator is the sum of each eigenvalue times its respective probability,

$$\langle J_1\rangle = -\frac{\hbar}{2}\mathscr{P}_{j_1'=-1/2} + \frac{\hbar}{2}\mathscr{P}_{j_1'=1/2} = \frac{\hbar}{2}\left(-\sin^2\frac{\omega_L t}{2} + \cos^2\frac{\omega_L t}{2}\right) = \frac{\hbar}{2}\cos\omega_L t, \tag{5.5.65}$$

which agrees with (5.5.52a) if j_1 is taken to be 1/2 in that equation.

5.6 Magnetic Resonance

Magnetic resonance is an important phenomenon described by a Hamiltonian that is an explicit function of time. While the magnetic resonance experiment discussed in this section can be used to determine the Landé g-factor, there are also many other important applications of the magnetic resonance phenomenon. Magnetic resonance imaging is now used extensively in medicine to examine soft tissue much the way that x-rays have been used to view bones. Also, functional magnetic resonance imaging is used, for example, in scientific research to determine which parts of the brain are active when specific tasks are performed.

From the previous section, the Hamiltonian of a rotator with magnetic moment $\boldsymbol{\mu} = g(e/2m)\mathbf{J}$ in the presence of an external magnetic field $\mathbf{B}$ is

$$H_0 = \frac{\mathbf{J}^2}{2I} - \boldsymbol{\mu}\cdot\mathbf{B} = \frac{\mathbf{J}^2}{2I} - g(e/2m)\mathbf{J}\cdot\mathbf{B} \equiv H^0 - g(e/2m)\mathbf{J}\cdot\mathbf{B}\,. \tag{5.6.1}$$

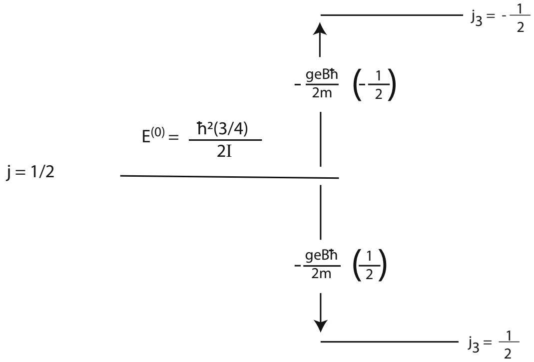


Fig. 5.5 Energy levels of a spin-1/2 rotator in a constant, uniform magnetic field. The Landé g-factor is assumed to be positive

If the direction of the 3 axis is chosen in the direction of **B**, the energy levels are

$$E_{j,j_3} = \frac{\hbar^2 j(j+1)}{2I} - \frac{geB\hbar}{2m} j_3 \,. \tag{5.6.2}$$

It is not crucial or even important that the quantum system placed in the magnetic field is a rotator. The Hamiltonian H^0 could just as well describe atoms, molecules or some other system provided that it satisfies two conditions: (1) The Hamiltonian H^0 must be rotationally invariant, implying that it commutes with J_i, $[H^0, J_i] = 0$. The first term in (5.6.2) would then be $E^0_{n,j}$ instead of $\hbar^2 j(j+1)/2I$, where n represents some additional quantum numbers. (2) The magnetic field **B** must be chosen so that the change in energy as a result of a spin flip is much smaller than $E^0_{n,j}$, implying that $E^0_{n,j}$ does not change during the experiment.

The energy levels for the spin $j = 1/2$ doublet are shown in Fig. 5.5. Before the magnetic field **B** is turned on, the energy $E^0_{n,j} = \hbar^2 j(j+1)/2I$ is degenerate in j_3. But, as a result of the interaction with the magnetic field, the degeneracy is removed because the energy depends on j_3. Since the value of j remains fixed (usually $j = 1/2$), if g is positive, energy levels with negative values of j_3 are shifted upward, and those with positive values are shifted downward.

If the mass m is known, g can be determined from the splitting of the energy levels. For example, taking $j = 1/2$ in (5.6.2),

$$g = \frac{2m}{eB\hbar}(E_{j=1/2,\, j_3=-1/2} - E_{j=1/2\,,j_3=1/2}) = \frac{2m}{eB\hbar}\Delta E \,. \tag{5.6.3}$$

The requirement that the energy gaps between the eigenvalues $E^0_{n,j}$ of the Hamiltonian H^0 be large in comparison with the energy splitting $g(e\hbar/2m)Bj_3$ is

$$|E^0_{n,j} - E^0_{n',j'}| \gg \Delta E \,. \tag{5.6.4}$$

The above condition can always be satisfied by choosing the magnetic field B to be sufficiently small. The value $E^0_{n,j}$ will then remain constant throughout the experiment, and the energy difference ΔE within a multiplet will be exceedingly small with a typical value $\Delta E \simeq 5 \times 10^{-6}$ eV. Thus the frequency of light emitted during a transition is $\nu = \Delta E/h \simeq 5 \times 10^{-6}$ eV/h $\simeq 10^9$ Hz, which is in the radio frequency range, while $E^0_{n,j} - E^0_{n,j'} \simeq 10^{-2}$ eV.

An effective method for determining the value of g for different systems is to measure the frequency of precession,

$$\omega_L = g\frac{eB}{2m}\,, \tag{5.6.5}$$

from which g follows immediately. Once g is known, the energy difference ΔE in (5.6.3) can also be calculated.

Expressed in terms of ω_L, the energy levels (5.6.2) can be written

$$E_{j,j_3} = \frac{\hbar^2 j(j+1)}{2I} - \hbar\omega_L j_3 \equiv E_j - \hbar\omega_L j_3\,, \tag{5.6.6}$$

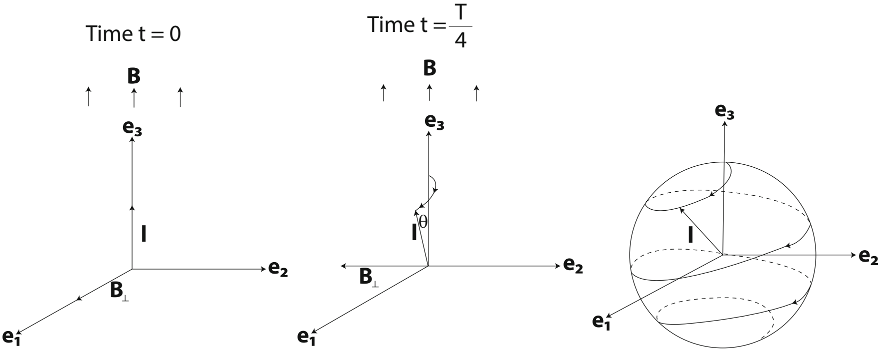
and the Hamiltonian (5.6.1) can be written as

$$H_0 = \frac{\mathbf{J}^2}{2I} - \omega_L J_3 \equiv H^0 - \omega_L J_3\,. \tag{5.6.7}$$

5.6.1 Hamiltonian for Magnetic Resonance

To determine the splitting between the two energy levels with $j = 1/2$ in Fig. 5.5 on the preceding page, a transition must be induced from the lower energy level to the higher level, which is accomplished by using a second magnetic field $B_\perp(t)$. Assuming g is positive as shown in Fig. 5.5 on the facing page, this transition would require the spin to flip from the spin-up position $j_3 = 1/2$ to the spin-down position $j_3 = -1/2$. In the same way that the spin $\mathbf{J}(t)$ or magnetic dipole $\boldsymbol{\mu}(t) = g(e/2m)\mathbf{J}(t)$ precesses about the $\mathbf{e}_3$ direction due to the magnetic field $\mathbf{B}$ in the 3-direction, a field $B_\perp(t)$ causes $\boldsymbol{\mu}(t)$ to precess about the direction of $B_\perp(t)$. Thus the spin can be flipped from the spin-up to the spin-down position by a second magnetic field denoted $\mathbf{B}_\perp$ that points in a direction $\mathbf{e}$ perpendicular to the $\mathbf{e}_3$-direction. After the magnetic dipole moment has precessed 180° about $\mathbf{B}_\perp$ from the up position, the spin has flipped to the down position.

If the magnetic dipoles $\boldsymbol{\mu}(t)$ were not already precessing about $\mathbf{e}_3$ as a result of $\mathbf{B}$, $\mathbf{B}_\perp$ could just be taken to point in the $\mathbf{e}_1$-direction. But since the magnetic dipole is already precessing about the $\mathbf{e}_3$ axis, $\mathbf{B}_\perp$ must also rotate about the $\mathbf{e}_3$ axis with a frequency ω_L so that it keeps in step with the rotating $\boldsymbol{\mu}(t)$ that rotates in a clockwise direction with an angular velocity ω_L as given in (5.6.5) . If instead $\mathbf{B}_\perp$



(a) (b) (c)

Fig. 5.6 Precession of a classical magnetic dipole caused by a constant, uniform magnetic field **B** in the 3-direction and a magnetic field B with a constant magnitude that rotates in the x-y plane

rotated in the opposite direction or with a frequency very different from ω_L, then its effect would just average to zero or produce minor high frequency oscillations in the precession about $\mathbf{e}_3$.

To better understand the effect of the field $\mathbf{B}_\perp$, the motion of a classical magnetic dipole moment is depicted in Fig. 5.6.

As shown in Fig. 5.6a, at $t = 0$, $\mathbf{B}_\perp$ is assumed to be in the $\mathbf{e}_1$-direction, and the angular momentum and $\boldsymbol{\mu}(t = 0)$ are in the $\mathbf{e}_3$-direction. The effect of the magnetic field $\mathbf{B}_\perp$, is that the angular momentum begins precessing about the $\mathbf{e}_1$ axis, moving toward the positive $\mathbf{e}_2$ axis. One quarter of a period later, as shown in Fig. 5.6b, the angular momentum vector $\mathbf{J}$ has precessed 90° about the $\mathbf{e}_3$ axis because of $\mathbf{B}$, which is in the $\mathbf{e}_3$-direction. Additionally, $\mathbf{J}$ has precessed an angle θ downward from the 3 axis because of $\mathbf{p}$. If the torque that $\mathbf{B}_\perp$ exerts on the magnetic dipole moment is to continue to cause precession from the positive $\mathbf{e}_3$ axis toward the negative $\mathbf{e}_3$ axis, during the quarter of a period $\mathbf{B}_\perp$ must also rotate 90° from the positive $\mathbf{e}_1$ axis to the negative $\mathbf{e}_2$ axis. Classically the angular momentum will spiral from the positive $\mathbf{e}_3$ axis to the negative $\mathbf{e}_3$ axis. The path of the tip of the classical angular momentum vector is shown in Fig. 5.6c. From this classical discussion the additional field $\mathbf{B}_\perp$ that would cause the spin to precess from the positive $\mathbf{e}_3$ axis to the negative $\mathbf{e}_3$ axis must rotate in a clockwise direction (as viewed from positive e_3 axis) in the 1–2 plane at a frequency ω (which must be approximately equal to ω_L). Thus $\mathbf{B}_\perp(t)$ is given by

$$\mathbf{B}_\perp(t) = B_\perp \mathbf{e}\,(t) = B_\perp(\mathbf{e}_1 \cos \omega t - \mathbf{e}_2 \sin \omega t)\,, \tag{5.6.8}$$

where $\mathbf{B}_\perp(t)$ has arbitrarily been chosen to point in the e_1 direction when $t = 0$.

The Hamiltonian, which determines the dynamics (time evolution) of the system, is

$$H(t) = \frac{\mathbf{J}^2}{2I} - g\frac{e}{2m}\mathbf{B} \cdot \mathbf{J} - g\frac{e}{2m}\mathbf{B}_\perp(t) \cdot \mathbf{J} = H_0 + H_1(t)\,. \tag{5.6.9}$$

The second and third terms in the above equation are of the same form with each giving the energy of the magnetic dipole as a result of one of the two magnetic fields. Expressing g in terms of ω_L using (5.6.5), (5.6.9) can be rewritten in the convenient form

$$H(t) = \frac{\mathbf{J}^2}{2I} - \omega_L J_3 - \frac{\omega_L}{B} B_\perp \mathbf{e}(t) \cdot \mathbf{J}\,. \tag{5.6.10}$$

Using the explicit expression $\mathbf{e}(t)$ that appears in (5.6.8),

$$H(t) = \frac{\mathbf{J}^2}{2I} - \omega_L J_3 - \omega_L \frac{B_\perp}{B}(J_1 \cos \omega t - J_2 \sin \omega t)\,. \tag{5.6.11}$$

The last term can be rewritten by setting $\theta = -\omega t$ in (5.5.29), which yields the form of H that will serve as the starting point for the solution:

$$H(t) = \frac{\mathbf{J}^2}{2I} - \omega_L J_3 - \omega_L \frac{B_\perp}{B} e^{i\omega t J_3/\hbar} J_1 e^{-i\omega t J_3/\hbar} = H_0 + H_1(t)\,. \tag{5.6.12}$$

The Hamiltonian (5.6.12) is the Hamiltonian for a magnetic dipole in the presence of a constant magnetic field $\mathbf{B}$ pointing in the $\mathbf{e}_3$ direction plus a term proportional to $B_\perp/B$ that accounts for the interaction with the rotating magnetic field. The experimental arrangement is such that the magnitude $B_\perp$ of the rotating magnetic field is much smaller than the magnitude B of the static field, $B_\perp \ll B$. Therefore, the motion is dominated by the uniform precession about the $\mathbf{e}_3$ axis. In the Schrödinger picture this is described, according to (5.5.37), by $\phi(t) = e^{-i[\hbar j(j+1)t/2I]}\, e^{i\omega_L J_3 t/\hbar}\phi(0) = e^{-iH_0 t/\hbar}\phi(0)$.

5.6.2 *Magnetic Resonance in the Schrödinger Picture*

Magnetic resonance is discussed in the Schrödinger picture using the Hamiltonian (5.6.12). Because this Hamiltonian *is an explicit function of time*, $\frac{\partial H(t)}{\partial t} \neq 0$ and the Schrödinger equation (5.3.9) is *not* simply solved by (5.3.3) with $U(t)$ as given by (5.3.8), or by (5.5.37). That is, $\phi(t)$ is *not* given by $e^{-iH(t)\,t/\hbar}\phi(0)$. With a little insight, however, the problem can be reformulated so that it is readily solved. To set the stage for the solution and further explain the choice of a rotating magnetic field, the case $B_\perp = 0$ discussed in the previous section is first reconsidered.

Classically, *if the experiment is viewed from a frame of reference that rotates with an angular frequency ω_L about the 3 axis in a clockwise direction as viewed from above*, the direction of the angular momentum vector does not change. If a detector is put in this moving frame, the expectation value of any component of angular momentum remains constant in time. So in the rotating frame there does not seem to be a magnetic field $\mathbf{B}$ acting on the magnetic dipole moment $\boldsymbol{\mu}$. If the additional field

$\mathbf{B}_\perp(t) = B_\perp \mathbf{e}(t)$, which is given in (5.6.8) and rotates with the angular frequency $\omega = \omega_L$, is applied, in the rotating frame the magnetic dipole moment would only experience a force resulting from $B_\perp$ and would precess about $\mathbf{e}(t)$. If $B_\perp$ rotates at $\omega = \omega_L$ in the 1–2 plane, from the point of view of the rotating frame, $\mathbf{B}_\perp$ is constant and $\mathbf{B}$ always points in the 3-direction. In this frame the Hamiltonian is time-independent! The strategy, then, is to first solve the Schrödinger equation for the state vector in the rotating frame and then transform the solution to the laboratory frame.

In the laboratory frame the state $\phi(t)$ obeys the Schrödinger equation with the Hamiltonian given by (5.6.12),

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = \left(\frac{\mathbf{J}^2}{2I} - \omega_L J_3 - \omega_L \frac{B_\perp}{B} e^{i\omega t\, J_3/\hbar} J_1 e^{-i\omega t\, J_3/\hbar} \right) \phi(t)\,. \tag{5.6.13}$$

Because $e^{i\omega t J_3/\hbar}$ commutes with J^2 and J_3, (5.6.13) can be rewritten as

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = e^{i\omega t J_3/\hbar} \left(\frac{\mathbf{J}^2}{2I} - \omega_L J_3 - \omega_L \frac{B_\perp}{B} J_1 \right) e^{-i\omega t J_3/\hbar} \phi(t)\,. \tag{5.6.14}$$

Multiplying (5.6.14) by the unitary operator $e^{-i\omega t J_3/\hbar}$ and using

$$e^{-i\omega t J_3/\hbar} i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-i\omega t J_3/\hbar}\phi(t) \right) - \omega J_3 e^{-i\omega t J_3/\hbar}\phi(t)\,, \tag{5.6.15}$$

the Schrödinger equation becomes

$$\begin{aligned} i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-i\omega t J_3/\hbar}\phi(t) \right) - \omega J_3 e^{-i\omega t J_3/\hbar}\phi(t) \\ = \left(\frac{\mathbf{J}^2}{2I} - \omega_L J_3 - \omega_L \frac{B_\perp}{B} J_1 \right) e^{-i\omega t J_3/\hbar}\phi(t)\,. \end{aligned} \tag{5.6.16}$$

Just looking at the mathematical form of (5.6.16) without any physical insight, the state $\phi(t)$ always appears in the combination $e^{-i\omega t J_3/\hbar}\phi(t)$. Consequently (5.6.16) is a differential equation for the vector $\phi(t)_R$ defined by

$$\phi(t)_R \equiv e^{-i\omega t J_3/\hbar}\phi(t)\,. \tag{5.6.17}$$

When written in terms of $\phi(t)_R$, (5.6.16) becomes

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\phi(t)_R = \left(\frac{\mathbf{J}^2}{2I} + (\omega - \omega_L) J_3 - \omega_L \frac{B_\perp}{B} J_1 \right) \phi(t)_R\,. \tag{5.6.18}$$

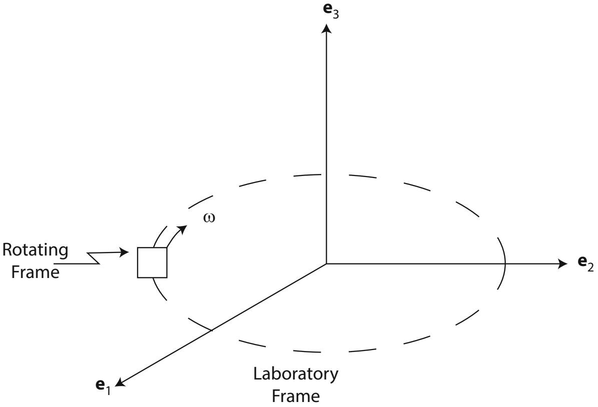


Fig. 5.7 A frame rotating about the 3 axis in a clockwise direction as viewed from above

The above equation is the Schrödinger equation for a new state vector $\phi(t)_R$ with the time-independent Hamiltonian

$$H_R = \frac{\mathbf{J}^2}{2I} + (\omega - \omega_L)J_3 - \omega_L \frac{B_\perp}{B} J_1 \, , \tag{5.6.19}$$

implying that the solution to (5.6.18) is

$$\begin{aligned}\phi(t)_R = e^{-iH_R t/\hbar}\phi(0)_R &= e^{-i[(\mathbf{J}^2/2I)+(\omega-\omega_L)J_3-\omega_L(B_\perp/B)J_1]t/\hbar}\phi(0)_R \\ &= e^{-i[H^0+(\omega-\omega_L)J_3-\omega_L(B_\perp/B)J_1]t/\hbar}\phi(0)_R \, .\end{aligned} \tag{5.6.20}$$

The above result could have been anticipated. When the magnetic field $\mathbf{B}_\perp(t)$, which rotates about the 3 axis with an angular velocity ω in a clockwise direction, was introduced, a time-dependent interaction term had to be added to the Hamiltonian. But in a frame that rotates precisely as $\mathbf{B}_\perp$ rotates, $\mathbf{B}_\perp$ is stationary and the Hamiltonian is time-independent. Thus $\phi(t)_R$ is just the state in the rotating frame.

In Fig. 5.7 the frame is shown rotating in a clockwise direction about the 3 axis with an angular velocity ω. From an apparatus in the rotating frame, any state in the laboratory frame of reference is rotating about the 3 axis in a counterclockwise direction with an angular velocity ω. After a time t, from the perspective of the rotating frame, a state in the lab frame will have rotated an angle ωt about the 3 axis in a counterclockwise direction. According to (5.5.22) the operator that effects this rotation is $e^{-i\omega t J_3/\hbar}$,

$$\phi(t)_{\text{rotating frame}} = e^{-i\omega t\, J_3/\hbar}\phi(t) \, , \tag{5.6.21}$$

where $\phi(t)$ is the state in the lab frame. Comparing (5.6.17) and (5.6.21), $\phi(t)_R$, which was introduced by examining the Schrödinger equation from a mathematical point of view, is indeed the state in the rotating frame, $\phi(t)_R = \phi(t)_{\text{rotating frame}}$. Regardless of whether mathematical or physical insight is used, the same conclusion is reached: the Hamiltonian is time-independent in the rotating frame.

Since any of these vectors has total angular momentum j, the first term in the exponent of (5.6.20) contributes only an overall phase $e^{-i[\hbar j(j+1)]/2I]t}$. Thus as can be seen from (5.6.20), in the rotating frame the time evolution of $\phi(t)_R$ resulting from the two magnetic fields $\mathbf{B}$ and $\mathbf{B}_\perp(t)$ is a rotation about $\mathbf{e}_1$ caused by $e^{-i[-\omega_L(B_\perp/B)J_1]t/\hbar}$ and a rotation about $\mathbf{e}_3$ caused by $e^{-i(\omega-\omega_L)J_3 t/\hbar}$. Using the same procedure employed in Example 5.5.4 on page 282, these two rotations can be expressed as a single rotation about a unit vector $\mathbf{n}$,

$$\mathbf{n} = \frac{\left(-\omega_L \frac{B_\perp}{B}\mathbf{e}_1 + (\omega-\omega_L)\mathbf{e}_3\right)}{\sqrt{\left(\omega_L \frac{B_\perp}{B}\right)^2 + (\omega-\omega_L)^2}}\,. \tag{5.6.22}$$

The component of the angular momentum along $\mathbf{n}$ is

$$\mathbf{n}\cdot\mathbf{J} = \frac{\frac{-\omega_L B_\perp}{B}\mathbf{e}_1\cdot\mathbf{J} + (\omega-\omega_L)\mathbf{e}_3\cdot\mathbf{J}}{\sqrt{\left(\omega_L \frac{B_\perp}{B}\right)^2 + (\omega-\omega_L)^2}} = \frac{-\omega_L \frac{B_\perp}{B} J_1 + (\omega-\omega_L) J_3}{\sqrt{\left(\omega_L \frac{B_\perp}{B}\right)^2 + (\omega-\omega_L)^2}}\,. \tag{5.6.23}$$

Using the above result, (5.6.20) can be written as

$$\begin{aligned}\phi(t)_R &= e^{-iH^0 t/\hbar} e^{-i\sqrt{\omega_L(B_\perp/B])^2+(\omega-\omega_L)^2}\,\mathbf{n}\cdot\mathbf{J}]t/\hbar}\phi(0)_R \\ &= e^{-iH^0 t/\hbar} e^{-i\,\Omega\,\mathbf{n}\cdot\mathbf{J}\,t/\hbar}\phi(0)_R\,,\end{aligned} \tag{5.6.24}$$

where Ω is defined by

$$\Omega = \sqrt{\left(\omega_L \frac{B_\perp}{B}\right)^2 + (\omega-\omega_L)^2}\,. \tag{5.6.25a}$$

From the discussion in the previous section, the exponential $e^{-i\,\Omega\,\mathbf{n}\cdot\mathbf{J}\,t/\hbar}$ describes a counterclockwise rotation about $\mathbf{n}$ by an angle

$$\sqrt{\left(\omega_L \frac{B_\perp}{B}\right)^2 + (\omega-\omega_L)^2}\;\; t \equiv \Omega t\,. \tag{5.6.25b}$$

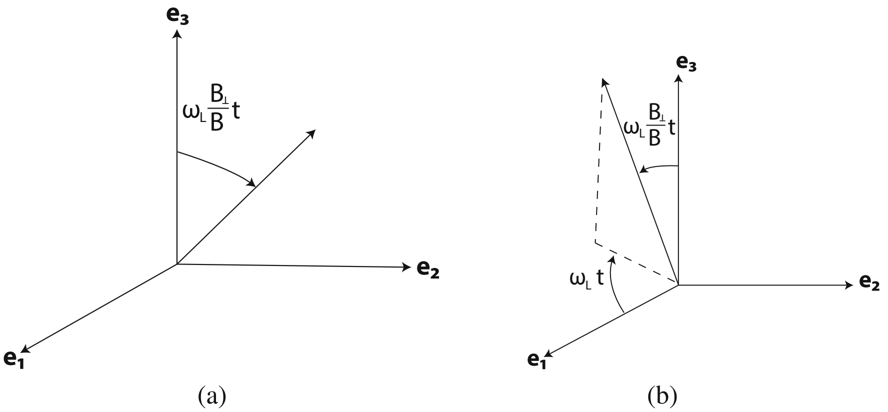


Fig. 5.8 Rotations required to obtain $\phi(t)$ from $\phi(0)$ in a magnetic resonance experiment when $\omega = \omega_L$. (**a**) Rotation by an angle $(\omega_L B_L/B)\,t$ about the $\mathbf{e}_1$ axis (**b**) Rotation by an angle $-\omega_L t$ about the $\mathbf{e}_3$ axis

Once the state $\phi(t)_R$ is known in the rotating frame, (5.6.21) immediately gives the state $\phi(t)$ in the laboratory frame,

$$\begin{aligned}\phi(t) &= e^{i\omega t J_3/\hbar}\phi(t)_R\,, \\ &= e^{i\omega t J_3/\hbar}e^{-iH^0 t/\hbar}e^{-i\,\Omega\,\mathbf{n}\cdot\mathbf{J}\,t/\hbar}\phi(0)_R\,. \end{aligned} \tag{5.6.26}$$

Noting from (5.6.21) that at $t = 0$ the state in the rotating frame coincides with that in the laboratory frame, $\phi(0)_R = \phi(0)$ so (5.6.26) can be expressed as

$$\phi(t) = e^{-iH^0 t/\hbar}e^{-i(-\omega t)J_3/\hbar}e^{-i\Omega\mathbf{n}\cdot\mathbf{J}t/\hbar}\phi(0)\,, \tag{5.6.27}$$

which is the result for any rotationally invariant Hamiltonian H^0. The above equation shows that the state $\phi(t)$ is obtained by first rotating the state $\phi(0)$ about $\mathbf{n}$ by an angle Ωt and then rotating the resulting state about $\mathbf{e}_3$ by an angle $-\omega t$. If $\omega = \omega_L$, then from (5.6.22) $\mathbf{n} = -\mathbf{e}_1$. For this case it then follows from (5.6.27) that the state first rotates about the $-\mathbf{e}_1$ axis by an angle $\Omega\,t = (\omega_L B_\perp/B)\,t$ and then rotates about the $\mathbf{e}_3$ axis by an angle $-\omega_L t$. These rotations are shown in Fig. 5.8. For the case that $B_\perp$ is zero, $\mathbf{n} = 0$ and (5.6.27) becomes the result (5.5.37), which is a state precessing clockwise about the $\mathbf{e}_3$ axis at the Larmor frequency ω_L. The second term in (5.6.11), which occurs because of the magnetic field in the 3 direction, causes the energy level $E_{n,j}$ $(= \hbar^2\, j(j+1)/2I$ for the rotator) to split into the two sublevels shown in Fig. 5.5 on page 284.

The probability is now calculated for obtaining the value $j_3 = -1/2$ in a measurement of J_3 at time t if, at time $t = 0$, the system is in a state $\phi(0)$ with $j_3 = 1/2$. That is, the initial state is one for which the probability of obtaining a value $j_3 = 1/2$ for the observable J_3 is unity. Thus $|\langle\phi(0)|j = 1/2,\ j_3 = 1/2\rangle|^2 =$

1 and $|\langle\phi(0)|j = 1/2,\ j_3 = -1/2\rangle|^2 = 0$, implying that, except for a possible phase, the state vector at time $t = 0$ is given by

$$\phi(0) = |j = 1/2,\ j_3 = 1/2\rangle\,. \tag{5.6.28}$$

The probability that the spin has flipped at a time t is the probability for obtaining the value $j_3 = -1/2$ in a measurement of J_3 in the state $\phi(t)$. Mathematically this probability is the expectation value of the projection operator

$$\Lambda_{-1/2} = |j = 1/2,\ j_3 = -1/2\rangle\langle j = 1/2,\ j_3 = -1/2|$$

in the state $\phi(t)$:

$$\mathscr{P}_{\phi(t)}(\Lambda_{-1/2}) = |\langle j = 1/2, j_3 = -1/2|\phi(t)\rangle|^2\,. \tag{5.6.29}$$

Using (5.6.27) and (5.6.28),

$$\begin{aligned}&\mathscr{P}_{\phi(t)}(\Lambda_{-1/2}) = \\ &= |\langle j = 1/2,\ j_3 = -1/2|e^{-i(\mathbf{J}^2/2I)t/\hbar}e^{-i(-\omega t)J_3/\hbar}e^{-i\Omega t\mathbf{n}\cdot\mathbf{J}/\hbar}|j = 1/2,\ j_3 = 1/2\rangle|^2\,.\end{aligned} \tag{5.6.30}$$

Recalling that $\mathbf{J}^2$ and J_3 are hermitian and letting $\mathbf{J}^2$ and then J_3 act to the left, the first two exponential terms only yield phase factors that don't contribute to an absolute value. The matrix element of the third exponential,

$$\mathscr{P}_{\phi(t)}\langle\Lambda_{-1/2}) = |\langle 1/2,\ -1/2|e^{-i\Omega t\mathbf{n}\cdot\hat{\mathbf{J}}/\hbar}|1/2,\ 1/2\rangle|^2\,, \tag{5.6.31}$$

can be simplified using (5.5.60) to yield

$$\mathscr{P}_{\phi(t)}\langle\Lambda_{-1/2}) = |\langle 1/2,\ -1/2|\cos\frac{\Omega t}{2} - \frac{2i}{\hbar}\,\mathbf{n}\cdot\mathbf{J}\sin\frac{\Omega t}{2}|1/2,\ 1/2\rangle|^2\,. \tag{5.6.32}$$

Inserting the explicit expression (5.6.23) for $\mathbf{J}$ into (5.6.32),

$$\begin{aligned}&\mathscr{P}_{\phi(t)}(\Lambda_{-1/2}) \\ &= \left|\langle 1/2,\ -1/2|\cos\frac{\Omega t}{2} - \frac{2i}{\hbar}\frac{\left(-\omega_L\frac{B_\perp}{B}J_1 + (\omega-\omega_L)J_3\right)}{\sqrt{\left(\omega\frac{B_\perp}{B}\right)^2 + (\omega-\omega_L)^2}}\ \sin\frac{\Omega t}{2}|1/2,\ 1/2\rangle\right|^2\,.\end{aligned} \tag{5.6.33}$$

From (2.3.37), (2.3.19), and (2.3.43), the matrix elements in (5.6.33) are found to have the following values:

$$\begin{aligned}
&\langle j=1/2, j_3=-1/2|j=1/2, j_3=1/2\rangle = 0\,,\\
&\langle j=1/2, j_3=-1/2|J_3|j=1/2, j_3=1/2\rangle = 0\,,\\
&\langle j=1/2, j_3=-1/2|J_1|j=1/2, j_3=1/2\rangle = \hbar/2\,.
\end{aligned}$$

Thus for a transition from the lower level of Fig. 5.5 when $j_3 = +1/2$ to the upper level with $j_3 = -1/2$ that occurs when energy is absorbed, the probability for spin flip is

$$\mathscr{P}_{\phi(t)}(\Lambda_{-1/2}) = \frac{\left(\omega_L \frac{B_\perp}{B}\right)^2}{\left(\omega_L \frac{B_\perp}{B}\right)^2 + (\omega-\omega_L)^2} \sin^2\frac{\Omega t}{2}\,, \tag{5.6.34}$$

where Ω is given in (5.6.25a).

In (5.6.34) the probability is a product of two functions. As expected the periodic function $\sin^2(\Omega\, t/2)$ has a first maxima when the state has rotated half a revolution, $\Omega\, t = \pi$, and then has further maxima when the state has rotated $\Omega\, t = 3\pi, 5\pi \ldots$. The spin-flip probability is zero when the state has not rotated, $\Omega\, t = 0$, or when it has rotated an integer number of revolutions, $\Omega\, t = 2\pi, 4\pi, \ldots$.

The probability (5.6.34) also depends on the function $f(\omega)$ where, according to (5.6.34),

$$f(\omega) = \frac{\left(\frac{\gamma}{2}\right)^2}{(\omega-\omega_L)^2 + \left(\frac{\gamma}{2}\right)^2}\,, \qquad \frac{\gamma}{2} \equiv \omega_L \frac{B_\perp}{B}\,. \tag{5.6.35}$$

The function $f(\omega)$ is characterized by two parameters, the location of its maximum value at $\omega = \omega_L$ and by γ, its width at half the maximum value. The function is called a Lorentzian or a Breit-Wigner probability with resonance parameters ω_L and γ. Formulas of the form (5.6.35) occur in all resonance phenomena and are important functions in quantum physics. The probability of spin flip is plotted as a function of ω in Fig. 5.9 on the following page at the time $t = \pi/\Omega$.

If the frequency ω of the rotation of the magnetic field $\mathbf{B}_\perp(t)$ about the 3 axis as given in (5.6.8) is very different from ω_L, the maxima at $t = \pi/\Omega$, etc. are shallow as shown in Fig. 5.10a. But as ω approaches ω_L, the function $f(\omega)$ approaches unity, and the probabilities at $t = t_{\max} = \pi/\Omega$, etc. also become unity, implying with certainty that the spin has flipped as shown in Fig. 5.10b.

By choosing the initial state with spin up ($j_3 = +1/2$), the case has been considered for which the quantum system with magnetic moment $\boldsymbol{\mu} = g(e/2m)\mathbf{J}$ is initially in the lower of the two energy levels in Fig. 5.5 on page 284. The case where the spin is initially down could instead have been chosen. The probability that the spin has flipped up at time t, and energy has been removed, can also be shown

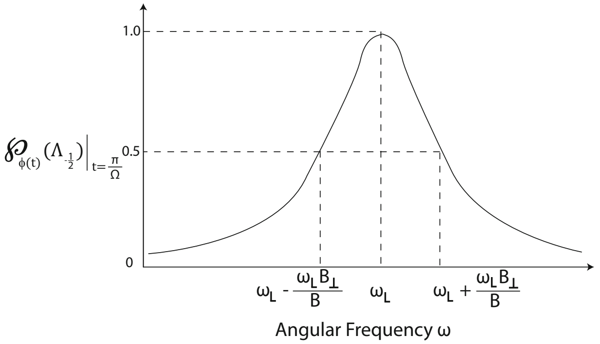
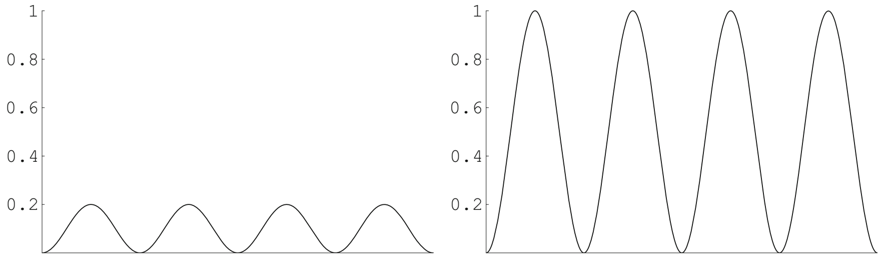


Fig. 5.9 Probability of a spin flip as a function of the angular frequency ω of the rotating magnetic field $\mathbf{B}_\perp$. The time has been chosen such that $\Omega t = \pi$



(a) (b)

Fig. 5.10 Transition probability between the spin-up and spin-down states caused by a rotating magnetic field $\mathbf{B}_\perp(t)$. Far from the resonance (**a**) the transition probability is small, and at resonance (**b**) there exist times when the transition probability is unity even when the field $\mathbf{B}_\perp(t)$ is small

to be given by (5.6.34) as can be readily checked. (See Problem 5.22.)

$$\mathscr{P}_{\phi(t)}(\Lambda_{1/2}) = \frac{\left(\omega_L \frac{B_\perp}{B}\right)^2}{\left(\omega_L \frac{B_\perp}{B}\right)^2 + (\omega - \omega_L)^2} \sin^2 \frac{\Omega t}{2}\,, \tag{5.6.36}$$

In deriving (5.6.34) under the influence of the additional rotating—and therefore time-dependent—magnetic field $\mathbf{p}(t)$, the system makes a transition from the lower to the upper energy level, $\phi(0) = |1/2, 1/2\rangle \to |1/2, -1/2\rangle$. Therefore, energy $\Delta E = ge\hbar B/2m$ is absorbed in this transition. On the other hand, for the case (5.6.36) the transition is from the upper to the lower energy level, $\phi(0) = |1/2, -1/2\rangle \to |1/2, 1/2\rangle$, and energy is removed from the quantum system. It is not obvious how to prepare initial states with spin down, $\phi(0) = |1/2, -1/2\rangle$, or with spin up, $\phi(0) = |1/2, 1/2\rangle$.

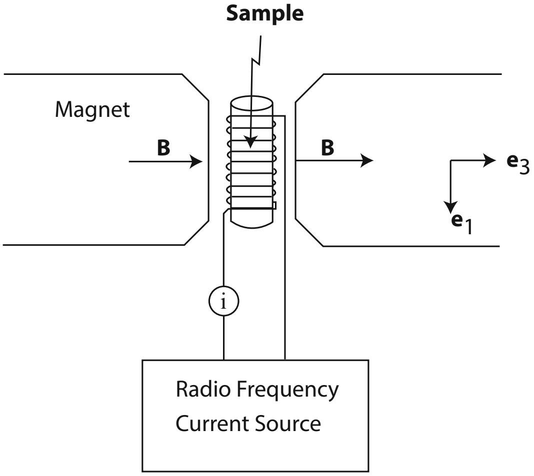


Fig. 5.11 Setup for a magnetic resonance experiment

5.6.3 *Magnetic Resonance Experiments*

The experimental setup for a magnetic resonance experiment is sketched in Fig. 5.11. The sample (water or glycerine for the measurement of the proton magnetic moment because H_2O molecules have zero magnetic moment except for that due to the protons) is placed in a field $\mathbf{B}$ that points in the $\mathbf{e}_3$ direction. The sample is, in addition, surrounded by a coil powered by a radio-frequency current source that produces a small field $\mathbf{B}_\perp$ that oscillates in the $\mathbf{e}_1$ direction with a variable frequency ω. Energy is absorbed by the sample while making the spin-flip transition, causing a detectable change in the current in the coils. When the energy supplied to the coils is a maximum, the frequency ω of the rotating field is identical to the Larmor frequency ω_L. Since the value of $\mathbf{B}$ is also known, the gyromagnetic ratio $g = 2m\omega_L/eB$ can be calculated. (In practice it is easier to change $\mathbf{B}$ than ω. Therefore, $\omega_L = eB/2m$ is changed by changing $\mathbf{B}$ until resonance occurs, which implies $\omega_L = \omega$.)

The careful reader will have noted that the time-dependent magnetic field described in the experimental setup does not actually rotate in the laboratory about the 3 axis with a frequency ω as required by (5.6.8). Instead, the time-dependent magnetic field used in the experiment is given by

$$\begin{aligned}\mathbf{B} &= 2B_\perp(\mathbf{e}_1 \cos \omega t) \\ &= B_\perp(\mathbf{e}_1 \cos \omega t - \mathbf{e}_2 \sin \omega t) + B_\perp(\mathbf{e}_1 \cos \omega t + \mathbf{e}_2 \sin \omega t)\,. \end{aligned} \tag{5.6.37}$$

The first field in (5.6.37) is the desired field, which rotates about the 3 axis in a clockwise direction when viewed from the positive 3 axis. The second field rotates in the opposite direction and can therefore never be in step with the rotating state. As a consequence, the second field in (5.6.37) never causes resonance behavior (See

Problem 5.25.) and only introduces high-frequency wiggles in the precession of the magnetic dipole moments about the 3 axis. These high frequency wiggles have no net effect on the average motion of the magnetic dipole moments.

The molecules in the sample of water used in the apparatus depicted in Fig. 5.11 on the previous page are in thermal equilibrium and are not in one of the pure states $\phi(0) = |1/2, \pm 1/2\rangle$ before the magnetic field is turned on. In thermal equilibrium the state of the protons in the water sample is a mixed state given by the Gibbs distribution that will be discussed in the next section. As will be shown, there are slightly more protons in the lower energy level with spin up than in the upper energy level with spin down. Therefore, when the rotating magnetic field $\mathbf{B}_\perp(t)$ is switched on, there are slightly more protons in the lower energy state making transitions to the upper energy state than there are protons in the upper energy state making transitions to the lower energy state. Consequently, a net, detectable energy must be supplied by the magnet coils to the sample, and this energy is a maximum at $\omega = \omega_L$. The maximum in energy is thus detected by measuring the current supplied to the coil and determining when it is a maximum. Once the resonance frequency ω_L and $|\mathbf{B}|$ have been determined to very high precision, from (5.6.5) the Landé g-factor can be determined to a very high accuracy. For water the magnetic dipole moment is twice that of a proton and g = 2 × (2.792847351 ± 0.000000028).

5.7 Gibbs Distribution

When the constant magnetic field $\mathbf{B} = B\,\mathbf{e}_3$ is introduced, the energy levels $E_j = \hbar^2 j(j+1)/2I$ split into sublevels, $\hbar^2 j(j+1)/2I - g(e\hbar/2m)Bj_3$, as shown for the case $j = 1/2$ in Fig. 5.5 on page 284. Before the constant magnetic field is introduced, the state of the ensemble of protons is given by the projection operator ρ^0 onto the two-dimensional angular momentum subspace,

$$\rho^0 = \frac{1}{2}[\Lambda_{j=1/2} + \Lambda_{j=-1/2}]\,, = \frac{1}{2}[|1/2, j_3 = 1/2\rangle\langle 1/2, j_3$$
$$= 1/2| + |1/2, j_3 = -1/2\rangle\langle 1/2, j_3 = -1/2|]\,. \tag{5.7.1}$$

For the sample at thermal equilibrium, the state of the ensemble of protons is given by the Gibbs state that will now be constructed.

Classically if there are N_1 objects with energy E_1 at thermal equilibrium with N_2 objects with energy E_2, then according to the Gibbs distribution[7]

$$\frac{N_1}{N_2} = \frac{e^{-E_1/kT}}{e^{-E_2/kT}}\,, \tag{5.7.2}$$

---

[7] L. D. Landau and E. M. Lifshitz, *Statistical Physics 5* V.1 (3 ed.) Pergamon Press, Oxford, 1980).

where k is Boltzmann's constant and T is the absolute temperature. Note that fewer objects are in the state with higher energy because the Boltzmann factor $e^{-E/kT}$ decreases as the energy E increases. From (5.7.2) the number of objects N_i in the ith state with energy E_i is proportional to $e^{-E_i/kT}$. That is,

$$N_i = Ce^{-E_i/kT} \,. \tag{5.7.3}$$

The proportionally constant C can be expressed in terms of the total number of objects $N = N_1 + N_2 + N_3 + \ldots$ in the system. Summing over the index i in (5.7.3),

$$\sum N_i = N = \sum_i Ce^{-E_i/kT} \,,$$

it follows that the constant C is given by

$$C = \frac{N}{\sum_j e^{-E_j/kT}} \,. \tag{5.7.4}$$

Using (5.7.3) and (5.7.4), for a collection with N objects, the probability that an object is in the ith state is

$$\frac{N_i}{N} = \frac{Ce^{-E_i/kT}}{N} = \frac{e^{-E_i/kT}}{\sum_j e^{-E_j/kT}} \,. \tag{5.7.5}$$

If the objects with energy E_i are quantum systems in the i^{th} energy eigenstate $|E_i\rangle$ satisfying $H|E_i\rangle = E_i|E_i\rangle$, then from (1.3.7) and (1.3.12) the statistical operator ρ is given by

$$\rho = \sum_i \frac{N_i}{N} |E_i\rangle\langle E_i| \,. \tag{5.7.6}$$

Example 5.7.1 Using the fact that the expectation value of an observable A is $\text{Tr}(\rho A)$, show that the probability $\mathscr{P}_\rho(A)$ of detecting energy E_i in the state ρ is N_i/N.

Solution The operator with an expectation value equal to the probability of detecting energy E_i in the state ρ, is $A = |E_i\rangle\langle E_i|$. Thus

$$\mathscr{P}_\rho(|E_i\rangle\langle E_i|) = \text{Tr}(\rho|E_i\rangle\langle E_i|) = \text{Tr}\sum_j \frac{N_j}{N} |E_j\rangle\langle E_j|E_i\rangle\langle E_i| \,.$$

Using the normalization condition $\langle E_j|E_i\rangle = \delta_{i,j}$ and taking the trace

$$\mathscr{P}_\rho(|E_i\rangle\langle E_i|) = \sum_k \sum_j \frac{N_j}{N}\langle E_k|E_j\rangle\delta_{i,j}\langle E_i|E_k\rangle = \frac{N_i}{N}\,.$$

Combining (5.7.5) and (5.7.6), the statistical operator ρ for a quantum system at thermal equilibrium is

$$\rho = \frac{\sum_i e^{-E_i/kT}}{\sum_j e^{-E_j/kT}}|E_i\rangle\langle E_i|\,. \tag{5.7.7}$$

Because $|E_i\rangle$ is an eigenstate of H with an eigenvalue E_i,

$$\rho = \frac{\sum_i e^{-H/kT}}{\sum_j \langle E_j|e^{-H/kT}|E_j\rangle}|E_i\rangle\langle E_i|\,. \tag{5.7.8}$$

The desired form for the statistical operator is obtained by recalling that

$$I = \sum_i |E_i\rangle\langle E_i|\,,$$

and noting that

$$\sum_j \langle E_j|e^{-H/kT}|E_j\rangle = \mathrm{Tr}(e^{-H/kT})\,.$$

Equation (5.7.8) can then be written as

$$\rho = \frac{e^{-H/kT}}{\mathrm{Tr}\left(e^{-H/kT}\right)}\,, \tag{5.7.9}$$

which is the statistical operator for a quantum system in thermal equilibrium.

For the ensemble of spin-1/2 magnetic dipole moments precessing in a constant magnetic field $\mathbf{B}$ along the $\mathbf{e}_3$ axis, this Gibbs state is now used to calculate the probability that a dipole is in a spin-down state, $j_3 = -1/2$, or in a spin-up state, $j_3 = 1/2$. Restricting to the subspace in which the angular momentum has the single eigenvalue $j = 1/2$, the notation is simplified by writing $|j, j_3\rangle \equiv |j_3\rangle$ where $j_3 = \pm 1/2$.

The denominator in (5.7.9) is calculated first using the explicit expression (5.5.34) for the energy.

$$\begin{aligned}\mathrm{Tr}\left(e^{-H/kT}\right) &= \sum_{j_3} e^{-E_{j_3}/kT} = e^{-E(j_3=-1/2)/kT} + e^{-E(j_3=1/2)/kT} \\ &= e^{-[3\hbar^2/8I-ge\hbar B/2m(-1/2)]/kT} + e^{-[3\hbar^2/8I-ge\hbar B/2m(1/2)]/kT} \\ &= e^{-3\hbar^2/8IkT}\left[e^{-ge\hbar B/4mkT} + e^{ge\hbar B/4mkT}\right].\end{aligned} \tag{5.7.10}$$

Substituting (5.7.10) into (5.7.9) and canceling the common factor $e^{-3\hbar^2/8IkT}$, the probability of spin down is

$$\mathscr{P}_\rho(\Lambda_{-1/2}) = \frac{e^{-ge\hbar B/4mkT}}{e^{-ge\hbar B/4mkT} + e^{ge\hbar B/4mkT}}\,. \tag{5.7.11}$$

Similarly, the probability of spin up is

$$\mathscr{P}_\rho(\Lambda_{+1/2}) = \frac{e^{ge\hbar B/4mkT}}{e^{-ge\hbar B/4mkT} + e^{gq\hbar B/4mkT}}\,. \tag{5.7.12}$$

To obtain an estimate of the relative magnitudes of the above two probabilities, their ratio is calculated for a proton ($g = 2.79$) at room temperature (293K) in a typical magnetic field $B = 1.0$ T, implying $ge\hbar B/4mkT = \omega_L(\hbar/2kT) = 1.74 \times 10^{-6}$. Therefore,

$$\frac{\mathscr{P}_\rho(\Lambda_{+1/2})}{\mathscr{P}_\rho(\Lambda_{-1/2})} = \frac{e^{ge\hbar B/4mkT}}{e^{-ge\hbar B/4mkT}} \cong 1 + 3.49 \times 10^{-6}\,. \tag{5.7.13}$$

As expected for the Gibbs state (5.7.9), the probability that the magnetic dipole is in the higher-energy, spin-down state is less than the probability that it is in the lower-energy, spin-up state. From (5.7.9) and (5.7.10), the density (statistical) operator for the spin-1/2 magnetic dipole moment in a constant magnetic field $\mathbf{B}$ along the $\mathbf{e}_3$ axis is

$$\rho(t=0)$$

$$= \frac{\left(e^{-ge\hbar B/4mkT}\,|j_3=-1/2\rangle\langle j_3=-1/2| + e^{ge\hbar B/4mkT}\,|j_3=1/2\rangle\langle j_3=1/2|\right)}{\left(e^{-gq\hbar B/4mkT} + e^{gq\hbar B/4mkT}\right)}\,, \tag{5.7.14a}$$

$$\equiv w_-|j_3=-1/2\rangle\langle j_3=-1/2| + w_+|j_3=1/2\rangle\langle j_3=1/2|\,. \tag{5.7.14b}$$

In thermal equilibrium there are more particles in the lower energy state with spin orientation $j_3 = 1/2$ so the sample has a minute excess magnetic moment parallel to the $\mathbf{e}_3$ axis.

Now consider the situation where the additional magnetic field $\mathbf{B}_\perp(t)$ is turned on at time $t = 0$, and the time development of the state $\rho(t)$, described by the Hamiltonian (5.6.11), is calculated. As a result of the time-dependent magnetic field, the state vector $|j_3 = 1/2\rangle$ evolves into the state vector $|j_3 = -1/2\rangle$ and vice versa. To obtain $\rho(t)$ from $\rho(t = 0)$, the Schrödinger state $\phi(t)$ is first expressed in terms of $\phi(0)$. From (5.6.27),

$$\phi(t) = e^{i\omega J_3 t/\hbar}\, e^{-i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\phi(0)\,, \tag{5.7.15}$$

where Ω is given by (5.6.25a). The Schrödinger states at time t that are in the spin down state $|j_3 = -1/2\rangle$ and spin up state $|j_3 = 1/2\rangle$ at $t = 0$, respectively, are denoted $\phi(t)_{\text{down}}$ and $\phi(t)_{\text{up}}$. From (5.7.15),

$$\phi(t)_{\text{down}} = e^{i\omega J_3 t/\hbar}\, e^{-i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\left|-\frac{1}{2}\right\rangle, \tag{5.7.16a}$$

$$\phi(t)_{\text{up}} = e^{i\omega J_3 t/\hbar}\, e^{-i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\left|\frac{1}{2}\right\rangle. \tag{5.7.16b}$$

At a later time t the Schrödinger states have evolved in time so that the density operator is given by

$$\begin{aligned}\rho(t) &= w_-|\phi(t)_{\text{down}}\rangle\langle\phi(t)_{\text{down}}| + w_+|\phi(t)_{\text{up}}\rangle\langle\phi(t)_{\text{up}}|\,,\\ &= w_- e^{i\omega J_3 t/\hbar}\, e^{-i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\left| -\frac{1}{2}\right\rangle\left\langle -\frac{1}{2}\right| e^{i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\, e^{-i\omega J_3 t/\hbar}\\ &\quad + w_+ e^{i\omega J_3 t/\hbar}\, e^{-i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\left|\frac{1}{2}\right\rangle\left\langle\frac{1}{2}\right| e^{i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\, e^{-i\omega J_3 t/\hbar}\,,\\ &= e^{i\omega J_3 t/\hbar}\, e^{-i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\rho(0) e^{i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\, e^{-i\omega J_3 t/\hbar}\,.\end{aligned} \tag{5.7.17}$$

Since the probability (5.6.34) for flipping the spin from up to down is the same as the probability (5.6.36) for flipping the spin from down to up, there would be no net energy absorption by the sample if both states were initially equally populated. As will now be shown, the excess population of the lower energy level described by the thermal equilibrium state (5.7.14) makes the measurement of ω_L possible.

In analogy with the calculation of (5.6.33) for $\mathscr{P}_{\phi_{\text{up}}(t)}(\Lambda_{-1/2})$ with $\phi_{\text{up}}(0) = |j_3 = 1/2\rangle$, the probability $\mathscr{P}_{\rho(t)}(\Lambda_{-1/2})$ is calculated for the angular momentum component to be $j_3 = -1/2$ in the state $\rho(t)$ that is in thermal equilibrium (5.7.14) at time $t = 0$. This probability is given by

$$\begin{aligned}\mathscr{P}_{\rho(t)}(\Lambda_{-1/2}) &= \mathrm{Tr}\left(\rho(t)|j_3 = -\frac{1}{2}\rangle(j_3 = -\frac{1}{2}|\right) \\ &= \sum_{j_3=-1/2}^{j_3=1/2} \langle j_3|\rho(t)| - \frac{1}{2}\rangle\langle - \frac{1}{2}|j_3\rangle = \langle - \frac{1}{2}|\rho(t)| - \frac{1}{2}\rangle\,. \qquad (5.7.18)\end{aligned}$$

Inserting $\rho(t)$ from (5.7.17) into (5.7.18) and then using the expression for $\rho(0)$ given in (5.7.14b),

$$\begin{aligned}\mathscr{P}_{\rho(t)}(\Lambda_{-1/2}) = \langle - \frac{1}{2}|e^{i\omega J_3 t/\hbar}\, e^{-i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\left[w_-| - \frac{1}{2}\rangle\langle - \frac{1}{2}| + w_+|\frac{1}{2}\rangle\langle\frac{1}{2}|\right] \\ \times\, e^{i[(\mathbf{J}^2/2I)+\Omega\mathbf{n}\cdot\mathbf{J}]t/\hbar}\, e^{-i\omega J_3 t/\hbar}| - \frac{1}{2}\rangle\,. \qquad (5.7.19)\end{aligned}$$

Since the operator $\mathbf{J}^2$ commutes with all other operators in (5.7.19) and always has the same value $j(j+1)\hbar^2 = 3\hbar^2/4$, it contributes only a phase that cancels out of the formula. The terms $e^{\pm i\omega J_3 t/\hbar}$ also yield phases that cancel. Thus,

$$\begin{aligned}\mathscr{P}_{\rho(t)}(\Lambda_{-1/2}) &= w_-\langle - \frac{1}{2}|e^{-i\Omega\mathbf{n}\cdot\mathbf{J}t/\hbar}| - \frac{1}{2}\rangle\langle - \frac{1}{2}|e^{i\Omega\mathbf{n}\cdot\mathbf{J}t/\hbar}| - \frac{1}{2}\rangle \\ &\quad + w_+\langle - \frac{1}{2}|e^{-i\Omega\mathbf{n}\cdot\mathbf{J}t/\hbar}| + \frac{1}{2}\rangle\langle + \frac{1}{2}|e^{i\Omega\mathbf{n}\cdot\mathbf{J}t/\hbar}| - \frac{1}{2}\rangle \\ &= w_-|\langle - \frac{1}{2}|e^{-i\Omega\mathbf{n}\cdot\mathbf{J}t/\hbar}| - \frac{1}{2}\rangle|^2 + w_+|\langle - \frac{1}{2}|e^{-i\Omega\mathbf{n}\cdot\mathbf{J}t/\hbar}|\frac{1}{2}\rangle|^2\,. \qquad (5.7.20)\end{aligned}$$

Since $j = 1/2$, the above equation can be simplified using (5.5.60),

$$\begin{aligned}\mathscr{P}_{\rho(t)}(\Lambda_{-1/2}) &= w_-|\langle - \frac{1}{2}|\left(\cos\frac{\Omega t}{2} - 2i\frac{\mathbf{n}\cdot\mathbf{J}}{\hbar}\sin\frac{\Omega t}{2}\right)| - \frac{1}{2}\rangle|^2 \\ &\quad + w_+|\langle - \frac{1}{2}|\left(\cos\frac{\Omega t}{2} - 2i\frac{\mathbf{n}\cdot\mathbf{J}}{\hbar}\sin\frac{\Omega t}{2}\right)|\frac{1}{2}\rangle|^2\,. \qquad (5.7.21)\end{aligned}$$

Using the expression for $\mathbf{n}\cdot\mathbf{J}$ of the dipole, which is given by (5.6.23), and the following values for matrix elements,

$$\begin{aligned}&\langle j_3 = -\frac{1}{2}|J_1|j_3 = -\frac{1}{2}\rangle = 0\,, &&\langle j_3 = -\frac{1}{2}|J_3|j_3 = -\frac{1}{2}\rangle = -\frac{\hbar}{2}\,, \\ &\langle j_3 = -\frac{1}{2}|J_1|j_3 = \frac{1}{2}\rangle = \frac{\hbar}{2}\,, &&(j_3 = -\frac{1}{2}|J_3|j_3 = \frac{1}{2}\rangle = 0\,,\end{aligned}$$

that are readily calculated from (2.3.19) and (2.3.43), (5.7.21) yields

$$\mathscr{P}_{\rho(t)}(\Lambda_{-1/2}) = w_- \left| \left(\cos\frac{\Omega t}{2} - 2i\frac{1}{2}\frac{(\omega-\omega_L)}{\sqrt{(\omega_L\frac{B_\perp}{B})^2+(\omega-\omega_L)^2}}\sin\frac{\Omega t}{2} \right) \right|^2$$

$$+ w_+ \left| \left(2i\frac{1}{2}\frac{(\omega_L\frac{B_\perp}{B})}{\sqrt{(\omega_L\frac{B_\perp}{B})^2+(\omega-\omega_L)^2}}\sin\frac{\Omega t}{2} \right) \right|^2 . \tag{5.7.22}$$

The probability of obtaining the value $j_3 = -1/2$ can be written in a more transparent form by introducing the quantity ϵ defined by $w_+ \equiv w_-(1+\epsilon)$. According to (5.7.13), at room temperature,

$$\frac{w_+}{w_-} = 1 + \epsilon = 1 + 3.49 \times 10^{-6} , \tag{5.7.23}$$

implying that $\epsilon = 3.49 \times 10^{-6}$. Then (5.7.22) can be rewritten in the desired form

$$\mathscr{P}_{\rho(t)}(\Lambda_{-1/2}) = w_- \left(1 + \epsilon \frac{(\omega_L\frac{B_\perp}{B})^2}{(\omega_L\frac{B_\perp}{B})^2+(\omega-\omega_L)^2}\sin^2\frac{\Omega t}{2} \right) . \tag{5.7.24}$$

The above formula is the analog of equation (5.6.34) that gives the probability of measuring the value $j_3 = -1/2$ at the time t if the initial state at $t = 0$ is the pure state $\rho(0) = |j_3 = 1/2\rangle\langle j_3 = 1/2|$. If all magnetic moments of the ensemble of hydrogen atoms in the water sample of the experiment in Fig. 5.11 on page 295 were in the lower state $|j_3 = 1/2\rangle$ of Fig. 5.5 on page 284 at $t = 0$, then the probability of measuring the magnetic moments in the upper state $|j_3 = -1/2\rangle$ of Fig. 5.5 is given by (5.6.34). In a realistic experiment such as that depicted in Fig. 5.11, the state of the magnetic moments of the hydrogen atoms is a mixture given by (5.7.14), describing an ensemble of a large number ($\approx N_{\text{Avogadro}}$) of magnetic moments distributed with weights w_+ and w_- over the two states $|j_3 = \pm 1/2\rangle$. After the state $\rho(t)$ in (5.7.17) has evolved as required by the Hamiltonian $H(t)$ in (5.6.9), the probability of measuring the value $j_3 = -1/2$ is given by (5.7.24). At time t the probability has increased by the second term in (5.7.24) over its value $\mathscr{P}_{\rho(0)}(\Lambda_{-1/2}) = w_-$ at time $t = 0$. The additional potential energy of the magnetic dipoles increases from a minimum at $t = 0$ to a maximum when $\Omega t/2 = \pi/2$ or $t = \pi/\Omega$. This additional energy is supplied by increased current to the transverse magnetic field coil. When the angular frequency ω with which $B_\perp(t)$ rotates in the plane perpendicular to $\mathbf{B}$ is varied, the probability (5.7.24) reaches its maximum when $\omega = \omega_L$. Determining the frequency at which maximum energy is supplied to the coil establishes the value of $\omega_L = geB/2m$. The Landé factor g can then be determined with high accuracy from ω_L.

If the state $\rho(t)$ of the magnetic moments of the hydrogen atoms in the water sample of the experiment in Fig. 5.11 on page 295 were not in thermal equilibrium (5.7.14), but instead were in an ensemble in which both states $|j_3 = \pm 1/2\rangle$ were equally populated, the state would be an ensemble in which $w_- = w_+ = 1/2$ or $\epsilon = 0$. Then from (5.7.24)

$$\mathscr{P}_{\rho(t)}(\Lambda_{-1/2}) = w_- = w_+ = \frac{1}{2}\,.$$

The probability of the transition would be constant in time and, more importantly, independent of the frequency ω. Only the term proportional to the minute factor ϵ in (5.7.24) leads to the typical Lorentzian resonance shape Fig. 5.9 on page 294 with a peak at the resonance frequency $\omega = \omega_L$ and a half-width $\gamma/2 = \omega_L B_\perp / B$.

5.8 Summary

Time development in quantum mechanics can described by the time-dependent Schrödinger (5.3.9) and by Heisenberg's equation of motion (5.4.3). In the Schrödinger picture all time dependence resides in the Schrödinger states $\phi_{\mathrm{S}}(t)$. If the Hamiltonian H is not an explicit function of time,

$$\phi_{\mathrm{S}}(t) = U(t)\phi_{\mathrm{S}}(0)\,, \quad \text{where} \quad U(t) = e^{-itH/\hbar}\,, \qquad -\infty < t < \infty\,.$$

$U(t)$ is the time-translation operator. The Schrödinger operators A_{S} do not change in time. In the Heisenberg picture all time dependence resides in the operators $A_{\mathrm{H}}(t)$. If the Hamiltonian is not an explicit function of time,

$$A_{\mathrm{H}}(t) = U^{\dagger}(t) A_{\mathrm{H}}(0) U(t)\,.$$

The Heisenberg states do not change in time.

If the Hamiltonian is not an explicit function of time, the Heisenberg and Schrödinger states and operators are related as follows:

$$\phi_{\mathrm{H}} = \phi_{\mathrm{S}}(0) = U^{\dagger}(t)\phi_{\mathrm{S}}(t)\,, \qquad A_{\mathrm{S}} = A_{\mathrm{H}}(0) = U(t) A_{\mathrm{H}}(t) U^{\dagger}(t)\,.$$

Eigenvectors of operators in the Schrödinger picture are constant in time and are given by $|a_n, t = 0\rangle$ while eigenvectors $|a_n, t\rangle$ of operators in the Heisenberg picture change in time,

$$|a_n, t\rangle = U^{\dagger}(t)|a_n, 0\rangle\,.$$

If the states and operators in any two pictures #1 and #2 are related by the relations

$$\phi_2 = U^\dagger \phi_1 \,, \quad A_2 = U^\dagger A_1 U \,,$$

where U is a unitary operator, the expectation value of an operator is the same in either picture.

When a classical spinning particle with mass m and charge q is placed in a constant magnetic field $\mathbf{B}$, the orbital angular momentum $\mathbf{l}$ of the particle satisfies

$$\frac{\mathrm{d}\mathbf{l}}{\mathrm{d}t} = \boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B} \,,$$

where $\boldsymbol{\mu} = (q/2m)\,\mathbf{l}$ is the classical magnetic moment of the spinning particle. If the z axis is chosen in the direction of $\mathbf{B}$, the solution to the above equation is

$$\begin{aligned} l_x(t) &= l_x(0)\, \cos\omega t + l_y(0)\, \sin\omega t \,, \\ l_y(t) &= -l_x(0)\, \sin\omega t + l_y(0)\, \cos\omega t \,, \\ l_z(t) &= l_z(0) \,, \end{aligned}$$

where the angular frequency $\omega = (q/2m)B$ is called the classical Larmor frequency. Classically the angular momentum precesses about the magnetic field at the classical Larmor frequency.

The Hamiltonian for a quantum particle with spin $\mathbf{J}$, moment of inertia I and magnetic moment $\boldsymbol{\mu} = g(q/2m)\mathbf{J}$ is

$$H = \frac{\mathbf{J}^2}{2I} - \boldsymbol{\mu} \cdot \mathbf{B} \,.$$

In the Schrödinger picture, if the system is in the initial state $|j, j_1\rangle$ at time $t = 0$, at a later time t the system is in the state $\phi(t)$,

$$\phi(t) = e^{-iHt/\hbar} |j, j_1\rangle = e^{-i[\hbar j(j+1)/2I]t} e^{i\omega_L t J_3/\hbar} |j, j_1\rangle \,,$$

where $\omega_L = g(q/2m)B$ is the Larmor frequency of a spinning particle with a Landé factor g. In the Schrödinger picture the operators J_i are time-independent, and the expectation value of an operator J_i at any time t is given by $\langle \phi_S(t) | J_i | \phi_S(t) \rangle$.

In the Heisenberg picture all time dependence resides in the operators J_i that obey Heisenberg's equation of motion,

$$\frac{\mathrm{d}J_i}{\mathrm{d}t} = -\frac{i}{\hbar}[J_i, H] + \frac{\partial J_i}{\partial t} \,.$$

Taking the constant, external magnetic field to point in the z direction, the solution to the above equation is

$$J_1(t) = J_1(0)\cos\omega_L t + J_2(0)\sin\omega_L t = e^{iHt}J_1(0)e^{-iHt}\,,$$

$$J_2(t) = -J_1(0)\sin\omega_L t + J_2(0)\cos\omega_L t = e^{iHt}J_2(0)e^{-iHt}\,,$$

$$J_3(t) = J_3(0) = e^{iHt}J_3(0)e^{-iHt}\,.$$

In the Heisenberg picture all time dependence resides in the operators. Thus if the system is in a state ϕ, the expectation value of the operator $J_i(t)$ at any time is $\langle\phi|J_i(t)|\phi\rangle$.

In an ensemble with N objects at thermal equilibrium, there are N_i objects with energy E_i. The classical probability of detecting an object with energy E_i is

$$\text{probability of detecting object with energy } E_i = \frac{N_i}{N} = \frac{e^{-E_i/kT}}{\sum_j e^{-E_j/kT}}\,.$$

The probability of detecting a particle with energy E_i decreases as E_i increases. From the above classical equation, it follows that the statistical operator ρ is given by

$$\rho = \frac{e^{-H/kT}}{\mathrm{Tr}\left(e^{-H/kT}\right)}\,.$$

Problems

For Sect. 5.3

5.1 Show that $U(t) = e^{-itH/\hbar}$ is a unitary operator, where the operator $H = H^\dagger$, by establishing the following:

(a) $U(0) = I\,,$
(b) $U^{-1}(t) = U(-t) = U^\dagger(t)$
(c) $U(t_1 + t_2) = U(t_1)Ut_2) = U(t_2)U(t_1)\,; \quad -\infty < t_1\,, t_2 < +\infty\,.$

5.2 Show that in the Schrödinger picture the density operator $\rho(t) = \sum_i \omega_i|\phi_i(t)\rangle\langle\phi_i(t)|$ satisfies the von Neuman equation (5.3.12)

For Sect. 5.4

5.3 Verify that the expectation value of an operator A in the pure state $\phi_S(t)$ is the same in the Schrödinger, Heisenberg, and interaction pictures.

5.4 By direct differentiation of (5.4.16), show that an operator $A_H(t)$ in the Heisenberg picture satisfies the Heisenberg equation of motion (5.4.3) provided $\frac{\partial H}{\partial t} = \frac{\partial A_S}{\partial t} = 0$.

5.5 By direct differentiation of (5.4.16), show that an operator $A_H(t)$ in the Heisenberg picture satisfies the Heisenberg equation of motion (5.4.3), where $\frac{\partial H}{\partial t} = 0$ but $\frac{\partial A_S}{\partial t} \neq 0$. The operator $\frac{\partial A_H(t)}{\partial t}$ is defined to be $e^{iHt/\hbar}\frac{\partial A_S}{\partial t}e^{-iHt/\hbar}$.

5.6 Verify the relation

$$e^{iB}Ae^{-iB} = A + i[B, A] + \frac{(i)^2}{2!}[B, [B, A]] + \frac{(i)^3}{3!}[B, [B, [B, A]]] + \dots .$$

Hint Consider a Taylor series expansion of

$$F(\lambda) = e^{i\lambda B}Ae^{-i\lambda B} = \sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\left(\frac{\partial^n F(\lambda)}{\partial\lambda^n}\right)\bigg|_{\lambda=0} .$$

By differentiating, it follows that

$$\frac{\partial F}{\partial\lambda} = e^{i\lambda B}i[B, A]e^{-i\lambda B}, \quad \frac{\partial^2 F}{\partial\lambda^2} = e^{i\lambda B}(i)^2[B, [B, A]]e^{-i\lambda B}, \quad \text{and so forth.}$$

5.7 Consider the harmonic oscillator in one-dimensional space described by the Hamiltonian $H = P^2/2m + KQ^2/2$ in the Schrödinger picture. Let $H_S^{(0)} = P^2/2m$ and $H_S' = KQ^2/2$.

(a) For the interaction picture calculate P_{ip} and Q_{ip} as linear functions of $P_S \equiv P$ and $Q_S \equiv Q$. (Hint: Use the identity in Problem 5.6.)

(b) Using your results from (a), calculate $H_{ip}^{(0)}$, H_{ip}', and H_{ip} as quadratic functions of P and Q.

5.8 By direct differentiation of $A_{ip}(t) = U^{(0)\dagger}(t)A_S U^{(0)}(t)$ show that

$$\frac{dA_{ip}}{dt} = \frac{i}{\hbar}\left[H_{ip}^{(0)}, A_{ip}\right] + \frac{\partial A_{ip}}{\partial t} ,$$

where $U^{(0)}(t) = e^{-iH_S^{(0)}t/\hbar}$ and $\frac{\partial A_{ip}}{\partial t} = U^{(0)\dagger}(t)\frac{\partial A_S}{\partial t}U^{(0)}(t)$. Assume that $H_S^{(0)}$ is not an explicit function of time.

5.9 By direct differentiation of $\phi_{\text{ip}}(t) = U^{(0)\dagger}(t)\phi_{\text{S}}(t)$ show that

$$i\hbar \frac{\mathrm{d}\phi_{\text{ip}}(t)}{\mathrm{d}t} = e^{iH^{(0)}t/\hbar} H'_{\text{S}} e^{-iH^{(0)}t/\hbar} \phi_{\text{ip}}(t) = H'_{\text{ip}} \phi_{\text{ip}}(t),$$

where $U^{(0)}(t) = e^{-iH_{\text{S}}^{(0)}t/\hbar}$. Assume that $H_{\text{S}}^{(0)}$ is not an explicit function of time.

For Sect. 5.5

5.10 Show that (5.5.9b) and (5.5.10) are a solution of (5.5.7) provided $c_3 = -c_1$, $c_2 = c_4$, and $\omega = geB/2m$.

5.11 Using (5.5.7), show that $\frac{\mathrm{d}(\mathbf{l}^2)}{\mathrm{d}t} = 0$.

5.12 Using (5.5.7), show that $\frac{\mathrm{d}(\mathbf{B}\cdot\mathbf{l})}{\mathrm{d}t} = 0$.

5.13 Show that $\langle j, j_1 | J_1 | j, j_1 \rangle = \hbar j_1$ and that $\langle j, j_1 | J_2 | j, j_1 \rangle = \langle j, j_1 | J_3 | j, j_1 \rangle = 0$. (Hint: Use results from Chap. 3 after making the substitution $z \to x$, $x \to y$, $y \to z$.)

5.14 Calculate $\langle j = 1/2, j_1 | J_3 | j = 1/2, j'_1 \rangle$ for the four possible combinations of values of j_1 and j'_1. (Hint: Use results from Chap. 2 after making the substitution $z \to x, x \to y, y \to z$.)

5.15 Calculate

$$\langle j = 1/2, j_3 = -1/2 | J_2 | j = 1/2, j_3 = 1/2 \rangle$$

and

$$\langle j = 1/2, j_1 = -1/2 | J_3 | j = 1/2, j_1 = 1/2 \rangle.$$

By drawing a coordinate system, explain why the above two matrix elements are equal.

5.16 By means of a diagram similar to Fig. 5.2 on page 272, explain why $e^{-i\theta J_3/\hbar} J_2 e^{i\theta J_3/\hbar}$ must equal $-J_1 \sin\theta + J_2 \cos\theta$. Verify this equality by using the identity (5.5.30).

5.17 Using the fact that in the Schrödinger picture operators are time-independent $A_{\text{S}}(t) = A_{\text{S}}(0)$, for the Hamiltonian in (5.5.34), calculate $J_{\text{H}}(t)$ using (5.4.16), which relates operators in the two pictures. Equation (5.5.29) and the corresponding relationship for J_2 given in the previous problem are useful.

5.18 Show that equations (5.5.48) can be written as the single vector equation

$$\mathbf{J}(t) = e^{-i\omega_L J_3(0)t/\hbar}\mathbf{J}(0)e^{i\omega_L J_3(0)t/\hbar}\,.$$

Would the above equation still be valid if $J_3(0)$ were replaced by $J_3(t)$ in the exponent?

5.19 Verify that $e^{-i\hbar j(j+1)t/2I}e^{i\pi J_3/(2\hbar)}|j, j_1\rangle$ is an eigenstate of J_2 with eigenvalue $-\hbar j_1$.

5.20 Show that for spin $j = 1/2$, $\phi(t)$ as given in (5.5.37) can be written as

$$\phi(t) = e^{-i\hbar j(j+1)t/2I}\left[\cos\frac{\omega_L t}{2} + \frac{2i}{\hbar}J_3 \sin\frac{\omega_L t}{2}\right]|j = 1/2, j_1\rangle.$$

Using the above result, calculate the expectation value of J_1 in the state $\phi(t)$ and verify that (5.5.39) is correct when $j = 1/2$.

For Sect. 5.6

5.21 The state vector in a frame rotating about the 3 axis with an angular velocity ω is related to the state vector in the laboratory frame by (5.6.17). Using a transformation of the form $H_R = U^\dagger HU$, calculate the Hamiltonian in the rotating frame if it is given by (5.6.12) in the laboratory frame.

5.22 Assume that the spin of the magnetic dipole moment is initially down. As a function of time, calculate the probability that the spin has flipped up.

5.23 Equation (5.6.34) gives the probability that the spin has flipped to the down position. By a direct calculation verify that the probability the spin has not flipped is

$$\mathscr{P}_{\phi(t)}(\Lambda_{1/2}) = \cos^2\frac{\Omega t}{2} + \frac{(\omega-\omega_L)^2}{\left(\omega_L\frac{B}{B}\right)^2 + (\omega-\omega_L)^2}\sin^2\frac{\Omega t}{2}.$$

From general principles what must be the value of the sum

$$\mathscr{P}_{\phi(t)}(\Lambda_{-1/2}) + \mathscr{P}_{\phi(t)}(\Lambda_{1/2})?$$

Verify that (5.6.34) and the expression for $\mathscr{P}_{\phi(t)}(\Lambda_{1/2})$ that you calculated satisfy this condition.

5.24 Calculate the expectation value of J_3 in the state $\phi(t)$ given by (5.6.27). Assume $\phi(0) = |j = 1/2, j_3 = 1/2\rangle$. Explain why

$$\langle\phi(t)|J_3|\phi(t)\rangle = \frac{\hbar}{2}\mathscr{P}_{\phi(t)}(\Lambda_{1/2}) - \frac{\hbar}{2}\mathscr{P}_{\phi(t)}(\Lambda_{-1/2}),$$

and check that your answer satisfies this condition.

5.25 Calculate the probability of spin flip for a field $\mathbf{B}_\perp = B_\perp(\mathbf{e}_1 \cos\omega t + \mathbf{e}_2 \sin\omega t)$ (which rotates in the counterclockwise direction when viewed from the positive 3 axis). Show that there is no resonance behavior.

For Sect. 5.7

5.26 An ensemble of harmonic oscillators in one-dimensional space is in thermal equilibrium at a temperature T. Find the probability that a harmonic oscillator has an energy $E_n = \hbar\omega(n + 1/2)$. Assume $E^{-\hbar\omega/kT} \ll 1$ and express the answer in terms of ω, T, and n.

Epilogue

The guiding principle in the presentation of quantum physics was the symmetry of the specially chosen systems, for which experimental data existed. This led us to the introduction of the basic concepts of quantum mechanics, including the notion of a state, an observable (linear operator in the vector space of the states) and the expectation value of an observable. By the same method, we have also discussed the time evolution of the physical systems. Our discussion introduced all mathematical concepts of quantum physics in an elementary way, while preserving the mathematical precision. We hope that a reader or a student, after studying our book will understand what quantum physics is, how it can be applied to the description of physical systems and above all, that he/she will understand the role which the symmetry plays in the description of the physical systems. This volume provides the conceptual basis of the forthcoming book, in which we will discuss further developments of the theory. A large part of our study will be devoted to the discussion of the scattering theory, theory of resonances and the description of unstable physical systems. We will also analyze the behavior of quantum systems in the slowly changing environment and the role of the quantum phase factors in such a description. As seen in this volume, our presentation will emphasize the simple description of complicated physical systems, using mathematically precise tools.

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9

Appendix: Mathematical Preliminaries

A.1 Introduction

A major new development in physics usually necessitates a corresponding development in mathematics. For example, differential and integral calculus were developed for classical mechanics to provide precise definitions for notions such as velocity and acceleration. It is true that special situations in mechanics can be treated without calculus, but the understanding remains vague without using calculus.

Quantum mechanics also has its own mathematical language, which was developed for the specific requirements of quantum physics—the physics of atoms, molecules, nuclei, and subatomic particles. The mathematical developments went hand-in-hand with the development of these areas of physics.

The mathematics of quantum mechanics uses vectors in linear, scalar-product spaces; linear operators; and algebras of operators in these spaces. Without the use of this mathematics, it is still possible to discuss some of the experimental data and to understand certain aspects of quantum mechanics. But the understanding is restricted in scope: a comprehensive understanding of the new ideas of quantum physics requires the use of its own mathematical language, which was discovered and developed for this very purpose.

Some of the key ideas of the mathematics of quantum mechanics as well as the notation are presented here without giving mathematical proofs. The treatment here is elementary, yet sufficiently detailed that it is possible to begin discussing quantum physics in its most general form with the mathematical tools presented here. Later, as required by the physics, additional mathematical concepts are introduced.

The mathematics that will be discussed here may be considered abstract as compared with differential operators or matrices. However, this is actually not the case. The one is as real as the other or, more precisely, as abstract as the other. A mathematical structure is a structure that exists in our minds. It is obtained by taking a set of mathematical objects and equipping this set with a structure by defining relations among these objects. Only familiarity makes some aspects of mathematics seem more real than others. Here in this introductory chapter, rather than treating

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9

the mathematics of quantum mechanics abstractly, each operation in a general linear space is motivated by first examining the corresponding operation in the familiar three-dimensional vector space. Also, when the properties of scalar-product spaces are discussed, each property is first shown to exist both for the scalar product in three-dimensional vector space and for the scalar product expressed as an integral.

The mathematical language of quantum mechanics was created so that quantum mechanics could be expressed in its general form. In 1926 P. Jordan and F. London started from the classical canonical transformations and recognized that these were coordinate transformations of a linear space. Physical quantities such as the intensity of radiation as an electron in an atom drops to a lower state were found to be represented by matrices. These matrices turned out to be matrices of operators in this linear space. First, Jordan and London considered only matrices and basis systems with discrete indices. The extension of the transformation theory to objects with continuous indices was done by Jordan and in particular by P.A.M. Dirac (1926–1927). Dirac's formalism was simple and beautiful but did not satisfy the requirement of mathematical rigor. The first rigorous mathematical formulation was given by D. Hilbert, L. Nordheim, and in particular by John von Neumann (1927) who associated the notions of quantum mechanical states and observables with vectors and operators, respectively, in the Hilbert space. Von Neumann's Hilbert space formulation could not accommodate objects with continuous indices and continuous eigenvalues. The mathematically rigorous formulation of quantum mechanics that includes the Dirac formalism, upon which our presentation here is based, was only possible after L. Schwartz (1950) had developed his distribution theory, and I.M. Gelfand and collaborators (1960) had introduced the rigged Hilbert space.

A.2 Linear, Scalar-Product Spaces

Linear spaces and linear operators are a generalization of certain aspects of three-dimensional space. The usual three-dimensional space consists of vectors that can be multiplied by real numbers and acted on by transformations or tensors. Mathematical objects such as vectors in three-dimensional space obey certain rules. To formulate the rules for a general, linear space, the rules from three-dimensional space are taken as the defining relations for a set of mathematical objects.

The linear spaces that are needed for quantum theory are, in general, not three-dimensional. They can have any dimension N, often infinite; the numbers are not real, but usually complex; the transformations are not orthogonal, but unitary; and the second rank tensors are not finite, but operators that can be represented by infinite matrices. In what follows, the rules for linear spaces are formulated in analogy with the usual rules for three-dimensional space.

Properties of the three -dimensional space

$\Re_3$

Under addition, two vectors $\mathbf{a}, \mathbf{b} \in \Re_3$ (i.e. $\mathbf{a}, \mathbf{b}$ in the space $\Re_3$) satisfy

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}. \tag{A.2.1a}$$

Addition is associative.

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}). \tag{A.2.2a}$$

There exists a zero vector with the property

$$\mathbf{0} + \mathbf{a} = \mathbf{a}. \tag{A.2.3a}$$

A vector can be multiplied with a real number b.

$$b(\mathbf{a}) = b\mathbf{a} \in \Re_3 \tag{A.2.4a}$$

Multiplication of a vector by real numbers a and b has the following properties:

$$a(b\mathbf{a}) = (ab)\mathbf{a} \tag{A.2.5a}$$

$$1\mathbf{a} = \mathbf{a} \tag{A.2.6a}$$

$$0\mathbf{a} = 0 \tag{A.2.7a}$$

Defining relations for the general linear space

Φ

The addition of two elements $\varphi, \psi \in \Phi$ is defined to satisfy

$$\varphi + \psi = \psi + \varphi. \tag{A.2.1b}$$

Addition is defined to be associative.

$$(\varphi + \psi) + \chi = \varphi + (\psi + \chi). \tag{A.2.2b}$$

There exists an element $0 \in \Phi$ with the property

$$0 + \psi = \psi\ . \tag{A.2.3b}$$

If $\psi \in \Phi$ and $b \in \mathbb{C}$ (b is a complex number), then

$$b(\psi) = b\psi \in \Phi\ . \tag{A.2.4b}$$

Multiplication of a vector by complex numbers a and b has, by definition, the following properties:

$$a(b\psi) = (ab)\psi \tag{A.2.5b}$$

$$1\psi = \psi \tag{A.2.6b}$$

$$0\psi = 0 \tag{A.2.7b}$$

On the left 0 is the number zero and on the right 0 is the element 0 of (A.2.7b).

Multiplication by real numbers satisfy

$$b(\mathbf{a} + \mathbf{b}) = b\mathbf{a} + b\mathbf{b}\,, \tag{A.2.8a}$$

$$(a + b)\mathbf{a} = a\mathbf{a} + b\mathbf{a}\,. \tag{A.2.9a}$$

The negative of a vector is defined by

$$(-1)\mathbf{a} = -\mathbf{a}\,. \tag{A.2.10a}$$

Multiplication by complex numbers satisfy

$$b(\varphi + \psi) = b\varphi + b\psi\ , \tag{A.2.8b}$$

$$(a + b)\psi = a\psi + b\psi\ . \tag{A.2.9b}$$

The negative of a vector is defined by

$$-1\psi = -\psi\ . \tag{A.2.10b}$$

Since **a**, **b**, and **c** are called vectors, the elements $\varphi, \psi \in \Phi$ are also called vectors. The set of mathematical objects φ, ψ, etc. that obey rules or axioms (A.2.1b)–(A.2.10b) is called a linear space; therefore, a linear space is defined by these rules alone. There are, of course, linear spaces whose objects have more properties than those stated above, but those additional properties are not necessary for them to be elements of a linear space.

One realization of an N-dimensional linear space would be by N-dimensional column matrices whose entries are complex. Another realization of a linear space is provided by complex, continuous, rapidly decreasing functions for which the functions themselves as well as all derivatives are square integrable. Because some people have spent more time studying functions as opposed to, say, matrices, one person may be more comfortable with one realization than another. Ultimately it is important to free oneself from all realizations and consider the linear space simply as a set of thought objects defined by (A.2.1b)–(A.2.10b). In physics the vectors of the linear space are realized by pure physical states. That is, these thought objects are used as mathematical images of physical states.

A linear space does not have enough structure to be of much use. To equip it with more structure, a scalar product is defined. Linear spaces with scalar products are called linear, scalar-product spaces, Euclidean spaces, or Pre-Hilbert spaces. In the usual three-dimensional space $\Re_3$, the scalar product of the vectors **a** and **b** is denoted by $\mathbf{a} \cdot \mathbf{b}$ and can be calculated using the formula

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z \,. \tag{A.2.11}$$

In a general linear space, the scalar product of two vectors φ and ψ will be denoted by either (φ, ψ) or $\langle \varphi | \psi \rangle$. Since there are some features of a scalar-product space that are not present in $\Re_3$, in addition to $\Re_3$, a realization of a scalar-product space by a space of "well behaved" functions will also be used. "Well behaved" means that all operations performed with the functions are well-defined. For such well-behaved functions $f(x)$ and $g(x)$, their scalar product is defined by

$$(f, g) \equiv \int_{-\infty}^{\infty} f^*(x) g(x) \mathrm{d}x \,. \tag{A.2.12}$$

Note that in contrast to (A.2.11), the scalar product (A.2.12) is, in general, complex. In formulating the rules for scalar-product spaces, the rule will first be examined in $\Re_3$ and then the corresponding rule will be considered for the complex, continuous functions just mentioned with a scalar product defined in (A.2.12). Finally, a general rule will be formulated. The scalar product is required to have the following properties:

Three-dimensional space $\Re_3$	Space of complex, continuous functions that, along with all derivatives are square integrable	General scalar-product space Φ
The scalar product of a vector with itself is positive definite, $\mathbf{a} \cdot \mathbf{a} \geq 0.$ (A.2.13a)	The scalar product of a function with itself is positive definite, $\int_{-\infty}^{\infty} f^*(x) f(x) \mathrm{d}x \geq 0.$ (A.2.13b)	In a linear space Φ, the scalar product of a vector with itself is positive definite. $(\psi, \psi) \geq 0.$ (A.2.13c)
$\mathbf{a} \cdot \mathbf{a} = 0$ iff (if and only if) $\mathbf{a} = 0$.	$\int_{-\infty}^{\infty} f^*(x) f(x) \mathrm{d}x = 0$ iff $f(x) = 0$.	For any $\psi \in \Phi$ $(\psi, \psi) = 0$ iff $\psi = 0$.
Since the scalar product is real, it trivially satisfies $\mathbf{a} \cdot \mathbf{b} = (\mathbf{b} \cdot \mathbf{a})^*.$ (A.2.14a)	The scalar product satisfies $(f, g) = \int_{-\infty}^{\infty} f^*(x) g(x) \mathrm{d}x = \left[\int_{-\infty}^{\infty} g^*(x) f(x) \mathrm{d}x\right]^* = (g, f)^*\,.$ (A.2.14b)	Any two vectors $\psi, \varphi \in \Phi$ must satisfy $(\psi, \varphi) = (\varphi, \psi)^*\,.$ (A.2.14c)
Multiplication by a real scalar a satisfies $a(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (a\mathbf{b}) = (a\mathbf{a}) \cdot \mathbf{b}\,.$ (A.2.15a)	Multiplication by a complex scalar satisfies $a(f, g) = \int_{-\infty}^{\infty} f^*(x)[a g(x)] \mathrm{d}x = (f, ag) = \int_{-\infty}^{\infty} [a^* f(x)]^* g(x) \mathrm{d}x = (a^* f, g)$ (A.2.15b)	For any $\psi, \varphi \in \Phi$ and any $a \in \mathbb{C}$, $a(\psi, \varphi) = (\psi, a\varphi) = (a^*\psi, \varphi)\,.$ (A.2.15c)

Note that while the convention (A.2.15c) for scalar products is standard in physics, it is not standard in the mathematical literature where one often finds $a(\psi, \varphi) = (\psi, a^*\varphi) = (a\varphi, \psi)$. That is, the scalar product in mathematical literature is often defined as the complex conjugate of the definition that is standard in the physics literature.

The scalar product of a sum of vectors is the sum of the scalar products.

$$(\mathbf{a}+\mathbf{b})\cdot\mathbf{c} = \mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c} \qquad \text{(A.2.16a)}$$

The scalar product of a sum satisfies

$$(f+g,h) = \int_{-\infty}^{\infty}[f(x)+g(x)]^{*}h(x)\mathrm{d}x = \int_{-\infty}^{\infty}f^{*}(x)h(x)\mathrm{d}x + \int_{-\infty}^{\infty}g^{*}(x)h(x)\mathrm{d}x\ . \qquad \text{(A.2.16b)}$$

For any $\psi, \varphi, \chi \in \Phi$ satisfies

$$(\psi+\varphi,\chi) = (\psi,\chi)+(\varphi,\chi) \qquad \text{(A.2.16c)}$$

The length or norm of a vector is

$$\|\mathbf{a}\| \equiv (\mathbf{a}\cdot\mathbf{a})^{\frac{1}{2}}. \qquad \text{(A.2.17a)}$$

The norm is defined by

$$\|f\| \equiv [\int_{-\infty}^{\infty}f^{*}(x)f(x)\mathrm{d}x]^{\frac{1}{2}}\ . \qquad \text{(A.2.17b)}$$

(φ,ψ) is called the scalar product of the vectors φ and ψ in the linear space Φ, and the space Φ is called a linear, scalar-product space or Euclidean space. In a scalar-product space the norm is defined by the scalar product

$$\|\psi\| \equiv (\psi,\psi)^{\frac{1}{2}}\ . \qquad \text{(A.2.17c)}$$

A vector is said to be normalized if $\|\psi\| = 1$.

Two vectors **a** and **b** are orthogonal if

$$\mathbf{a}\cdot\mathbf{b} = 0\ . \qquad \text{(A.2.18a)}$$

Two functions $f(x)$ and $g(x)$ are said to be orthogonal if

$$\int_{-\infty}^{\infty}f^{*}(x)g(x)\mathrm{d}x = 0\ . \qquad \text{(A.2.18b)}$$

Two vectors ψ and φ are defined to be orthogonal if

$$(\psi,\varphi) = 0\ . \qquad \text{(A.2.18c)}$$

Example A.2.1 Consider the following two functions on the interval $-\infty < x < \infty$:

$$f_1(x) = A_1 e^{-x^2/2} \qquad f_2(x) = A_2 x e^{-x^2/2}$$

Determine the constants $|A_1|$ and $|A_2|$ such that $f_1(x)$ and $f_2(x)$ are normalized. Are f_1(x) and f_2(x) orthogonal?

Solution The normalization conditions for the $f_i(x)$ are

$$1 = (f_i, f_i) = \int_{-\infty}^{\infty} f_i^{*}(x) f_i(x)\mathrm{d}x\ .$$

For $f_1(x)$, the above integral becomes

$$1 = |A_1|^2 \int_{-\infty}^{\infty} e^{-x^2} \mathrm{d}x = |A_1|^2 \sqrt{\pi},$$

where the integral was evaluated with the aid of a table of integrals. Thus,

$$|A_1| = \pi^{-\frac{1}{4}}.$$

Similarly,

$$1 = |A_2|^2 \int_{-\infty}^{\infty} x^2 e^{-x^2} \mathrm{d}x = \frac{\sqrt{\pi}}{2},$$

or

$$|A_2| = \sqrt{2}\pi^{-\frac{1}{4}}$$

To determine if $f_1(x)$ and $f_2(x)$ are orthogonal, their scalar product is calculated:

$$(f_1, f_2) = \int_{-\infty}^{\infty} f_1^*(x) f_2(x) \mathrm{d}x = A_1^* A_2 \int_{-\infty}^{\infty} x e^{-x^2} \mathrm{d}x = 0. \tag{A.2.19}$$

The above integral is zero from symmetry: By changing to the integration variable to $y = -x$, the integral is found to equal the negative of itself and is, therefore, zero. Since $(f_1, f_2) = 0$, the functions are orthogonal.

Example A.2.2 Using (A.2.14c), show that the relation $a(\psi, \varphi) = (\psi, a\varphi\ )$ implies the relation $a(\psi, \varphi) = (a^*\psi, \varphi\ )$ for $a \in \mathbb{C}$.

Solution Taking the complex number to be a^* instead of a, the first equality in (A.2.15c) immediately yields

$$a^*(\varphi, \psi) = (\varphi, a^*\psi).$$

Taking the complex conjugate of both sides of the above equation

$$a(\varphi, \psi)^* = (\varphi, a^*\psi)^*.$$

Using (A.2.14c) immediately yields the desired equality.

A.3 Linear Operators

The rules or axioms for linear operators in linear spaces are formulated here in analogy with operators in three-dimensional space.

Three-dimensional space $\Re_3$

Vectors in $\Re_3$ can be transformed into other vectors. One example is the rotation R, which rotates a vector $\mathbf{a}$ into a new vector $\mathbf{b} = R\mathbf{a}$. There are also other transformations such as the moment of inertia tensor I that transforms one vector into another according to $\mathbf{j} = I\boldsymbol{\omega}$. These transformations have the following properties:

$$R(\mathbf{a} + \mathbf{b}) = R\mathbf{a} + R\mathbf{b}, \tag{A.3.1a}$$

$$R(a\mathbf{b}) = a(R\mathbf{b}), \tag{A.3.2a}$$

where a is a real number.

Transformations (tensors) in $\Re_3$ can be added, multiplied by a real number, and multiplied by each other.

$$(R_1 + R_2)\mathbf{a} = R_1\mathbf{a} + R_2\mathbf{a}\,, \tag{A.3.3a}$$

$$(aR_1)\mathbf{a} = a(R_1\mathbf{a})\,, \tag{A.3.4a}$$

$$R_1 R_2\mathbf{a} = R_1(R_2\mathbf{a})\,. \tag{A.3.5a}$$

The zero transformation 0 and the identity transformation 1 exist in $\Re_3$ and are defined, respectively, by

$$0\mathbf{a} = 0\,, \tag{A.3.6a}$$

$$1\,\mathbf{a} = \mathbf{a}\,, \tag{A.3.7a}$$

for all $\mathbf{a} \in \Re_3$.

General scalar-product space Φ

In a linear, scalar-product space transformations or linear operators are defined as follows: A function or operator A that maps each vector $\psi \in \Phi$ into a vector $\varphi \in \Phi$,

$$\varphi = A(\psi) \equiv A\psi,$$

is called a linear operator if it obeys the rules (A.3.1b) and (A.3.2b) listed below. Thus, by definition, linear operators have the following properties:

$$A(\psi + \varphi) = A\psi + A\varphi, \tag{A.3.1b}$$

$$A(a\psi) = a(A\psi), \tag{A.3.2b}$$

where $a \in \mathbb{C}$.

Linear operators A and B can be added, multiplied by a complex number, and multiplied by each other. Then $A + B$, $a\,A$, and $A\,B$ are also linear operators that satisfy

$$(A + B)\psi = A\psi + B\psi \tag{A.3.3b}$$

$$(aA)\psi = a(A\psi) \tag{A.3.4b}$$

$$(A\,B)\psi = A(B\psi)\,. \tag{A.3.5b}$$

Operators of special interest are the zero operator 0 and the unit or identity operator $\mathbb{1}$ defined, respectively, by

$$0\,\psi = 0, \quad 0 \in \Phi\,, \tag{A.3.6b}$$

$$\mathbb{1}\,\psi = \psi\,, \tag{A.3.7b}$$

for all $\psi \in \Phi$.

For every hermitian operator R defined for all vectors in $\Re_3$, there exists a non-zero vector $\mathbf{b}$ such that

$$R\mathbf{b} = \lambda \mathbf{b}\,, \quad \lambda \in \mathbb{R}\,. \tag{A.3.8a}$$

The vector $\mathbf{b}$ is called an eigenvector of R and λ is a (real) eigenvalue of R.

For every operator R defined for all vectors, the transpose operator R^T has the following property: Writing the scalar product in components

$$\mathbf{a} \cdot (R\mathbf{b}) = \sum_{i=1}^{3}\sum_{j=1}^{3} a_i R_{ij} b_j$$

$$= \sum_{i=1}^{3}\sum_{j=1}^{3} a_i R^T_{ji} b_j$$

$$= \sum_{i=1}^{3}\sum_{j=1}^{3} (R^T_{ji} a_i) b_j = (R^T \mathbf{a}) \cdot \mathbf{b} \tag{A.3.9a}$$

If there exists a non-zero vector ψ such that

$$A\psi = \lambda \psi\,, \quad \lambda \in \mathbb{C}\,, \tag{A.3.8b}$$

then ψ is called an eigenvector of A and λ is called an eigenvalue of A.

A linear operator B in the space Φ is called the adjoint of the operator A if

$$(\psi, A\varphi) = (B\psi, \varphi)\,,$$

for all $\varphi, \psi \in \Phi$. The operator B is denoted $B = A^\dagger$. The operator A is called self-adjoint or hermitian[1] if

$$(\psi, A\varphi) = (A\psi, \varphi)\,, \tag{A.3.9b}$$

for all $\varphi, \psi \in \Phi$.

Example A.3.1 Find the adjoint of the operator $R = a\frac{d}{dx}$, $a \in \mathbb{C}$, for the space of complex, continuous functions which, along with all derivatives, are square integrable.

Solution Integrating by parts,

$$(f, R\,g) = \int_{-\infty}^{\infty} f^*(x)\, a \frac{\mathrm{d}g(x)}{\mathrm{d}x} \mathrm{d}x,$$

$$= af^*(x)g(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \left[a \frac{\mathrm{d}f^*(x)}{\mathrm{d}x}\right] g(x)\mathrm{d}x.$$

Since $f(\pm\infty) = g(\pm\infty) = 0$, the surface term vanishes. Therefore,

$$(f, R\,g) = \int_{-\infty}^{\infty} \left[-a^* \frac{\mathrm{d}f(x)}{\mathrm{d}x}\right]^* g(x)\mathrm{d}x.$$

From the definition of the adjoint operator, $R^\dagger = -a^* \frac{\mathrm{d}}{\mathrm{d}x}$.

---

[1] For unbounded operators, if (A.3.9b) is fulfilled in a mathematically vague way without specifying the domain, the operator is often said to be hermitian. When (A.3.9b) is defined precisely mathematically, the operator is said to be selfadjoint.

Example A.3.2 Show that the scalar product $(A\psi, \psi)$ is real if $A^\dagger = A$.

Solution Using (A.2.14c),

$$(A\,\psi, \psi)^* = (\psi, A\,\psi).$$

But from the definition (A.3.6b) for the adjoint of an operator,

$$(\psi, A\,\psi) = (A^\dagger\psi, \psi).$$

Combining the above two equations,

$$(A\,\psi, \psi)^* = (A^\dagger\psi, \psi).$$

Thus if $A^\dagger = A$, the complex conjugate of the scalar product $(A\psi, \psi)$ equals itself and is therefore real.

If ψ is an eigenvector of the operator A, an immediate consequence of Example A.3.2 is that all eigenvalues of a hermitian operator are real. For this reason, hermitian operators are often called real operators.

Example A.3.3 Show that any two eigenvectors ψ, ϕ of a hermitian operator A satisfying

$$A\,\psi = a_1\psi, \qquad A\,\phi = a_2\phi; \qquad a_1 \neq a_2,$$

have the property

$$(\psi, \phi) = 0.$$

Solution Since ϕ is an eigenstate of A,

$$(\psi, A\,\phi) = (\psi, a_2\phi) = a_2(\psi, \phi).$$

Using the Hermiticity of the operator A as given in (A.3.6b),

$$(\psi, A\,\phi) = (A\,\psi, \phi) = (a_1\psi, \phi) = a_1^*(\psi, \phi).$$

Remembering that the eigenvalues of a hermitian operator are real, subtracting the above two expressions for $(\psi, A\phi)$ yields

$$0 = (a_2 - a_1)(\psi, \phi).$$

Thus if $a_1 \neq a_2$, $(\psi, \phi) = 0$.

For any two vectors $\psi, \phi \in \Phi$ and an operator A, the scalar product of $A\,\psi$ with ϕ, namely $(\phi, A\,\psi)$, plays an important role in physics and is called the matrix element of the operator A between the vectors ϕ and ψ .

Example A.3.4 Calculate the four matrix elements $(f_i(x), xf_j(x))$ where $f_1(x)$ and $f_2(x)$, respectively, are the normalized functions

$$f_1(x) = \pi^{-\frac{1}{4}} e^{-x^2/2} \quad \text{and} \quad f_2(x) = \sqrt{2}\,\pi^{-\frac{1}{4}} x e^{-x^2/2}$$

first discussed in Example A.3.1 on page 321.

Solution From the definition of a matrix element,

$$(f_i(x), xf_j(x)) = \int_{-\infty}^{\infty} f_i^*(x) x f_j(x)\,\mathrm{d}x.$$

Since $f_1(x)$ and $f_2(x)$ are real,

$$(f_1(x), xf_2(x)) = (f_2(x), xf_1(x)) = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} e^{-x^2} x^2 \mathrm{d}x = \frac{1}{\sqrt{2}}.$$

Using the procedure mentioned in Example A.2.1 on page 318, the diagonal matrix elements are found to equal zero,

$$(f_1(x), xf_1(x)) = (f_2(x), xf_2(x)) = 0.$$

The results can be summarized by the single matrix

$$(f_i(x), xf_j(x)) = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix},$$

where the first index labels the row and the second labels the column of each matrix element.

All observables such as position, momentum, and energy are represented by linear operators in a linear, scalar-product space, and states are represented by vectors in this same space.

A.4 Basis Systems and Eigenvector Decompositions

A.4.1 Discrete Basis Vectors in Real, Three-Dimensional Space

Three basis vectors are introduced in the three-dimensional space $\Re_3$,

$$\mathbf{e}_i, \quad i = 1, 2, 3, \tag{A.4.1}$$

that span the space and are usually normalized to unity,

$$\mathbf{e}_i \cdot \mathbf{e}_i = 1. \tag{A.4.2}$$

These basis vectors are also usually chosen so that they are orthogonal to one another,

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0 \text{ if } i \neq j. \tag{A.4.3}$$

Instead of labeling the vectors by $i = 1, 2, 3$, they could also be denoted by $\mathbf{e}_1 = \mathbf{e}_x$, $\mathbf{e}_2 = \mathbf{e}_y$, $\mathbf{e}_3 = \mathbf{e}_z$. Relations (A.4.2) and (A.4.3) can be written as the single equation

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}, \quad i, j = 1, 2, 3, \tag{A.4.4}$$

where δ_{ij} is the Kronecker-δ. The set of orthogonal, normalized vectors is called an orthonormal set.

A basis system may be chosen arbitrarily provided it spans the space although, for a specific physical problem, one basis system may be much easier to work with than others. For example, for a rigid body with an inertia tensor I, it is helpful to choose the basis system such that the inertia tensor is diagonal. Therefore, the $\mathbf{e}_i$ are chosen such that

$$\mathbf{e}_i \cdot I \cdot \mathbf{e}_j = I(j)\,\delta_{ij} \text{ or } I \cdot \mathbf{e}_j = I(j)\,\mathbf{e}_j. \tag{A.4.5}$$

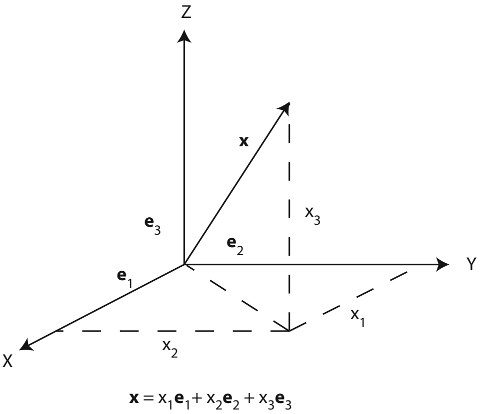
The $\mathbf{e}_j$ are the eigenvectors of the tensor I and the $I(j)$ are the eigenvalues.

In $\Re_3$ every vector $\mathbf{x}$ can be expanded in terms of the basis system,

$$\mathbf{x} = \sum_{i=1}^{3} x_i \mathbf{e}_i = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3. \tag{A.4.6}$$

The numbers x_i are the coordinates or components of $\mathbf{x}$ with respect to the $\mathbf{e_i}$. As shown in Fig. A.1 on the facing page, the x_1, x_2, and x_3 are, respectively, the x, y, and z-components of the vector $\mathbf{x}$.

Fig. A.1 The three-dimensional vector **x** expressed in terms of the normalized (unit) vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$



Taking the scalar product of both sides of (A.4.6) with $\mathbf{e}_j$,

$$\mathbf{e}_j \cdot \mathbf{x} = \sum_{i=1}^{3} \mathbf{e}_j \cdot \mathbf{e}_i x_i = \sum_{i=1}^{3} \delta_{ij} x_i = x_j, \tag{A.4.7}$$

where the orthogonality relation (A.4.4) has been used. Clearly the x_i determine the vector **x** uniquely. Using the expression for x_j in (A.4.7), (A.4.6) can be written in the form

$$\mathbf{x} = \sum_{i=1}^{3} \mathbf{e}_i (\mathbf{e}_i \cdot \mathbf{x}) . \tag{A.4.8}$$

Example A.4.1 Consider the two-dimensional vector **V** shown in Fig. A.2 on the next page, which has a magnitude $|\mathbf{V}|$, makes an angle θ with the $\tilde{x}$-axis, and makes an angle θ' with the $\mathbf{e_1}$-axis. From geometrical considerations, express **V** in terms of the unit vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ and then in terms of $\mathbf{e}_1$ and $\mathbf{e}_2$. Show that the second formula agrees with (A.4.7).

Solution The $\tilde{x}$ and $\tilde{y}$-components of **V** are clearly $|\mathbf{V}| \cos \theta$ and $|\mathbf{V}| \sin \theta$, respectively. Therefore,

$$\mathbf{V} = |\mathbf{V}| \cos\theta \; \tilde{\mathbf{x}} + |\mathbf{V}| \sin\theta \; \tilde{\mathbf{y}}.$$

Using similar logic, in terms of the orthonormal basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$,

$$\mathbf{V} = |\mathbf{V}| \cos \theta' \mathbf{e}_1 + |\mathbf{V}| \sin \theta' \mathbf{e}_2.$$

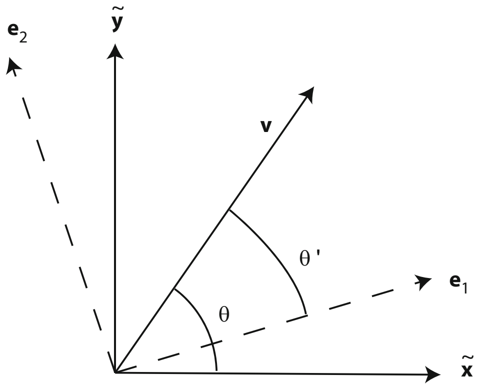


Fig. A.2 The two-dimensional vector **V** shown with respect to the $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$ basis and the $\mathbf{e}_1$, $\mathbf{e}_2$ basis

Now from the definition of the dot product,

$$\mathbf{e}_1 \cdot \mathbf{V} = |\mathbf{e}_1||\mathbf{V}| \cos\theta' = |\mathbf{V}| \cos\theta',$$

and

$$\mathbf{e}_2 \cdot \mathbf{V} = |\mathbf{e}_2||\mathbf{V}| \cos(90 - \theta'\rangle = |\mathbf{V}| \sin\theta'.$$

Combining the above three equations,

$$\mathbf{V} = \mathbf{e}_1(\mathbf{e}_1 \cdot \mathbf{V}) + \mathbf{e}_2(\mathbf{e}_2 \cdot \mathbf{V}),$$

which, for two-dimensional vectors **V**, agrees with (A.4.8).

In $\Re_3$, the scalar product of the vectors

$$\mathbf{x} = \sum_{i=1}^{3} \mathbf{e}_i x_i \quad \text{and} \quad \mathbf{y} = \sum_{j=1}^{3} \mathbf{e}_j y_j$$

is calculated as follows:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{3} \sum_{j=1}^{3} \mathbf{e}_i x_i \cdot \mathbf{e}_j y_j$$

Using (A.4.4),

$$\begin{aligned}\mathbf{x}\cdot\mathbf{y} &= \sum_{i=1}^{3}\sum_{j=1}^{3} x_i\ \delta_{ij}\ y_j\,,\\ \mathbf{x}\cdot\mathbf{y} &= \sum_{i=1}^{3} x_i\ y_i\,.\end{aligned} \tag{A.4.9}$$

The square of the norm (square of the length) of the vector **x** is given by

$$\|\mathbf{x}\|^2 = \mathbf{x}\cdot\mathbf{x} = \sum_{i=1}^{3} x_i\ x_i\,. \tag{A.4.10}$$

A.4.2 Discrete Basis Vectors in Infinite-Dimensional, Complex Space

In a linear, scalar-product space over the complex numbers, it is possible to introduce, in analogy to Fig. A.1 on page 325, a basis system in a three-dimensional, complex linear space for which the coordinates x_i of a vector ψ are in general complex numbers. Without any difficulties this three-dimensional space can be generalized to an N-dimensional space. To go from N dimensions to infinite dimensions is more difficult: the meaning of convergence of infinite sequences must be defined, which means that the topology of the linear space Φ must be defined. This can be done in many different ways, two of which result in the Hilbert space and the Schwartz space.

In an N-dimensional (or infinite-dimensional) space the basis vectors are denoted

$$e_n = |n) \quad \text{or} \quad e_n = |n\rangle \quad n = 1, 2, 3, \ldots N \quad \text{or} \quad \infty\,. \tag{A.4.11}$$

These vectors are again chosen to be orthonormal,

$$(e_i, e_j) = \langle i|j\rangle = \delta_{ij}\,. \tag{A.4.12}$$

In an N-dimensional (or infinite-dimensional), complex, linear, scalar-product space Φ, there exists an orthonormal basis system. That is, every vector $\psi \in \Phi$ can be expressed as

$$\psi = \sum_{n=1}^{N \text{ or } \infty} |e_n\rangle c_n = \sum_{n=1}^{N \text{ or } \infty} |e_n\rangle\langle e_n|\psi\rangle\,, \tag{A.4.13}$$

where the coordinates or components $c_n = \langle e_n|\psi\rangle$ are complex numbers. To make the expansions (A.4.13) mathematically rigorous, theorems are required that are not proved here. Instead the expansions are simply constructed in analogy to (A.4.8).

In analogy with the vectors $\mathbf{e}_i$ in the three dimensional space $\Re^3$ that satisfy (A.4.5), it is often convenient to choose the basis vectors $|e_n\rangle$ to be eigenvectors of a self-adjoint operator $A = A^\dagger$ that is of particular physical significance. The eigenvectors in the N-dimensional (or infinite-dimensional), linear, scalar-product space, are solutions of the eigenvalue equation

$$A|e_n\rangle = a_n|e_n\rangle, \quad |e_n\rangle \in \Phi, \quad n = 1, 2, 3 \ldots . \tag{A.4.14}$$

In Examples A.3.2 and A.3.3 on page 322 it was established that the eigenvalues a_i of a self-adjoint operator A are real and that two eigenvectors $|e_i\rangle$ and $|e_j\rangle$ with different eigenvalues $a_i \neq a_j$ are orthogonal. Thus it is possible to normalize[2] the eigenvalues in (A.4.14) such that (A.4.12) is fulfilled.

If $|e_i\rangle$ is a normalized eigenvector with eigenvalue a_i then

$$|e'_n\rangle = e^{i\omega}|e_n\rangle, \qquad \omega \in \Re\,, \tag{A.4.15}$$

is also a normalized eigenvector with the same eigenvalue a_i (See Problem A.10.) so the solutions of (A.4.14) are only determined up to a phase factor $e^{i\omega}$.

Since only the combination $|e_n\rangle\langle e_n|$ appears in (A.4.13) and

$$|e'_n\rangle\langle e'_n| = e^{i\omega}|e_n\rangle\langle e_n|e^{-i\omega} = |e_n\rangle\langle e_n|\,, \tag{A.4.16}$$

the choice of phase (A.4.15) is of no consequence. The important mathematical quantities are the projection operators $\Lambda_n = |e_n\rangle\langle e_n|$ that project onto orthogonal subspaces,

$$\Lambda_n\Lambda_m = \delta_{nm}\Lambda_n\,, \tag{A.4.17}$$

and are independent of the phase. The projection operators have the property that $\Lambda_n\mathbf{V} = |e_n\rangle\langle e_n|\mathbf{V}\rangle = v_n|e_n\rangle$, implying that the projection operator Λ_n projects out the component v_n of the vector $\mathbf{V}$ along the $\mathbf{e}_n$ axis as shown in Fig. A.3 on the facing page.

Two cases are now distinguished:

1. There is only one normalized eigenvector (up to a phase) for each eigenvalue of the chosen operator A. The eigenvalues are said to be nondegenerate, and the projection operator Λ_n projects onto a one-dimensional subspace.
2. There is more than one normalized eigenvector (up to the phase) for at least one eigenvalue of the operator A. The eigenvalues are said to be degenerate. A

---

[2] If $|e_i\rangle$ is not normalized and $\|\,|e_i\rangle\| \neq 0$, the new vector $|e'_i\rangle = |e_i\rangle/\|\,|e_i\rangle\|\,\}$ is a normalized eigenvector with the same eigenvalue.

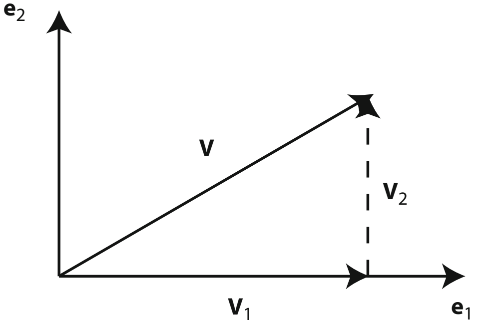


Fig. A.3 The two-dimensional vector **V** with components v_1 and v_2 along the respective basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$

discussion of Case 2 will be postponed until it is needed to describe physical problems.

Example A.4.2 Show that the functions

$$f_1(x, y) = xe^{-x^2/2}\, e^{-y^2/2} \quad \text{and} \quad f_2(x, y) = ye^{-x^2/2}\, e^{-y^2/2}$$

both satisfy

$$Af_i(x, y) = af_i(x, y), \quad i = 1, 2,$$

where the constant a is the eigenvalue of the operator

$$A = \frac{\mathrm{d}^2}{\mathrm{d}x^2} - x^2 + \frac{\mathrm{d}^2}{\mathrm{d}y^2} - y^2.$$

The coordinates x and y range from $-\infty$ to ∞ .

Solution Allowing the differential operator A to act on $f_1(x, y)$ and $f_2(x, y)$,

$$Af_1(x, y) = -4f_1(x, y), \quad Af_2(x, y) = -4f_2(x, y).$$

If the functions $f_1(x, y)$ and $f_2(x, y)$ are labeled only by their eigenvalue a,

$$f_{a=-4}(x, y) = f_1(x, y) \quad \text{and} \quad f_{a=-4}(x, y) = f_2(x, y).$$

Since there are two different eigenfunctions $f_{a=-4}(x, y)$, the eigenfunctions of A are not uniquely specified by the eigenvalue a.

For Case 1 the eigenvector $|e_n\rangle$ in (A.4.11) is uniquely determined (up to a phase) by its eigenvalue a_n. Therefore, the vector $|e_n\rangle$ can be labeled by the value a_n,

$$|e_n\rangle \equiv |a_n\rangle \quad n = 1, 2, 3, \dots N\,. \tag{A.4.18}$$

The eigenvalue equation (A.4.14) then becomes

$$A|a_n\rangle = a_n|a_n\rangle, \quad |a_n\rangle \in \Phi, \quad n = 1, 2, 3 \dots\,, \tag{A.4.19a}$$

$$\langle a_n|a_m\rangle = \delta_{nm}\,, \tag{A.4.19b}$$

and the basis vector expansion (A.4.13) becomes the eigenvector expansion

$$\psi = \sum_{n=1}^{N \text{ or } \infty} |a_n\rangle c_n = \sum_{n=1}^{N \text{ or } \infty} |a_n\rangle\langle a_n|\psi\rangle\,, \tag{A.4.20}$$

A complete system of eigenvectors possesses the property that every vector $\psi \in \Phi$ can be expanded in terms of the $|a_n\rangle$ according to the eigenvector expansion (A.4.20). Is it possible to find a complete system of eigenvectors $|a_n\rangle$ for every self-adjoint operator A in a linear scalar product space Φ? When Φ is finite-dimensional, the answer is "yes," and when Φ is infinite-dimensional, as is the case for the Hilbert space, the answer is "no."

A.4.3 Continuous Basis Systems of a Linear Space

Some properties of physical systems can be described by operators with a complete set of discrete eigenvectors $|a_n\rangle$, $n = 1, 2, 3, \dots$. The eigenvalues a_n represent the values observed in experiments on quantum physical systems. For example, if the operator A is the energy operator of the hydrogen atom, the eigenvalues a_n are the discrete energy values $E_n = -2\pi \mathrm{R}\hbar\mathrm{c}/n^2$, $n = 1, 2, 3, \dots$, where R is the Rydberg constant.

On the other hand it is impossible to describe all physical systems with operators that have discrete spectra. In addition to the discrete eigenvalues E_n corresponding to the electron-proton bound states of the hydrogen atom, an electron interacting with a proton also has continuous values of energy when there is no binding and the electron is scattered by the proton. The electron's motion corresponds to the hyperbolic orbits of the corresponding classical Kepler system. Another operator with a continuous eigenvalue spectrum is the momentum operator P of electrons in cathode rays. The operator P can have any of a continuous set of values depending on the accelerating potential. The position x must similarly be described by an operator with a continuous eigenvalue spectrum.

Thus in addition to the set of discrete eigenvectors $H|E_n\rangle = E_n|E_n\rangle$, eigenvectors of energy H, momentum P, position Q, as well as many other operators can have continuous eigenvalues:

$$H|E\rangle = E|E\rangle\,, \qquad 0 \le E < \infty\,, \tag{A.4.21a}$$

$$P|p\rangle = p|p\rangle\,, \qquad -\infty \le p \le +\infty\,, \tag{A.4.21b}$$

$$Q|x\rangle = x|x\rangle\,, \qquad -\infty \le x < +\infty \quad \text{or} \quad M \le x \le N\,. \tag{A.4.21c}$$

These are the eigenkets first introduced by Dirac. They are called generalized eigenvectors and, as in (A.4.21), are denoted by the symbol $|\ \rangle$ to indicate that the spectrum of the eigenvalue is continuous.

Since the eigenvalue x of the position operator Q is continuous, it is not possible to choose eigenvectors such that their (generalized) scalar product is a Kronecker-δ (A.4.19b) because a Kronecker-δ only involves discrete indices. In spite of this mathematical complication, it is possible to discuss generalized eigenvectors of operators with a continuous spectrum in analogy to the discrete case. The starting point is the eigenvector expansion, which was postulated by Dirac as the continuous analogue of (A.4.20).

There are spaces, an example of which is the Schwartz space Φ, for which any vector $\psi \in \Phi$ can be expanded in terms of the generalized eigenvectors $|x\rangle$ of Q according to

$$\psi = \int \mathrm{d}x\, |x\rangle\langle x|\psi\rangle \equiv \int \mathrm{d}x\, |x\rangle \psi(x)\,, \qquad \psi(x) \equiv \langle x|\psi\rangle\,. \tag{A.4.22a}$$

The above equation implies that

$$\langle\phi|\psi\rangle = \int \mathrm{d}x\, \langle\phi|x\rangle\langle x|\psi\rangle = \int \mathrm{d}x\, \phi(x)^*\psi(x)\,, \tag{A.4.22b}$$

where the integral extends over the continuous set of eigenvalues M< x< N where often $M = -\infty$ and $N = +\infty$.

To make the transition from (A.4.20) to (A.4.22a), the sum over the discrete variable n is replaced by a continuous sum (integral) over the continuous variable x, and the eigenvectors $|n\rangle$ are replaced by the generalized eigenvectors $|x\rangle$. Equation (A.4.22a) is called the generalized basis system expansion and was justified mathematically by the Nuclear Spectral Theorem.[3]

The coordinates or components $\psi(x) \equiv \langle x|\psi\rangle = (|x\rangle, |\psi\rangle)$ of the vector ψ with respect to the basis system $|x\rangle$ are the (generalized) "scalar product" between the vector ψ and the generalized eigenvector $|x\rangle$ and are complex numbers just as the coordinates $\langle a_n|\psi\rangle$ are complex numbers in the discrete case (A.4.20). For the continuous case the coordinates $\langle x|\psi\rangle$are continuous, well-behaved (infinitely

---

[3] After Dirac introduced (A.4.22a), approximately 30 years elapsed until Distribution Theory (L. Schwartz (1950–1951)) and the Rigged Hilbert Space (Gelfand et al., K. Maurin (1955–1959)) provided the mathematics for its justification.

differentiable, rapidly decreasing) functions of the continuous variable x whereas in (A.4.20) they are functions of the discrete variable a_n.

The progression from (A.4.8) to (A.4.22) can be taken as a justification of the mathematics underlying Dirac's eigenvector expansion (A.4.22); however, the mathematical proof is provided by the Nuclear Spectral Theorem and is required to make the mathematics rigorous.[4]

For the discrete case, taking the scalar product of ψ as given in (A.4.20) with the vector $|a_m\rangle$ gives,

$$(|a_m\rangle, \psi) \equiv \langle a_m|\psi\rangle = \sum_n \langle a_m|a_n\rangle\langle a_n|\psi\rangle\,, \tag{A.4.23}$$

which implies $\langle a_m|a_n\rangle = \delta_{mn} = \delta_{a_m\, a_n}$.

In analogy to (A.4.23) the (generalized) scalar product of ψ, as given in (A.4.22a), with the generalized eigenvector $|x'\rangle$ is written as

$$(|x'\rangle, \psi) \equiv \langle x'|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\langle x'|x\rangle\langle x|\psi\rangle. \tag{A.4.24}$$

The quantity $\langle x'|x\rangle$ is thus the analog of $\langle a_m|a_n\rangle$ in (A.4.23). Since $\langle a_m|a_n\rangle = \delta_{mn}$, where δ_{mn} is the Kronecker-δ, the generalized scalar product $\langle x'|x\rangle$ was also written by Dirac as

$$\langle x'|x\rangle = \delta(x - x), \tag{A.4.25}$$

where $\delta(x - x')$ is called the Dirac-δ functional, which is continuous. The Dirac-δ, $\delta(x' - x)$, is not a (locally integrable) function. Instead it is a new mathematical quantity called a functional that is defined by (A.4.24). Defining the function $\psi(x) \equiv \langle x|\psi\rangle$, (A.4.24) is then written as

$$\psi(x') = \int \mathrm{d}x\,\delta(x' - x)\,\psi(x). \tag{A.4.26}$$

This functional is defined as the mathematical object with the property that it gives the value of the well-behaved function at x'. That is, integrating from $-\infty < x < +\infty$ over the product of a Dirac-δ and a "well-behaved" function is the the mathematical procedure that maps $\psi(x)$ into $\psi(x')$, the value of the function at the position x'.

The generalized eigenvectors $|x\rangle$ are not vectors in the space Φ. The generalization $\langle x'|\psi\rangle$ is not an ordinary scalar product of a vector ψ with a vector $|x'\rangle$. The generalized scalar product can be made mathematically precise as an antilinear functional on the space Φ: $F_x(\psi) \equiv \langle\psi|x\rangle$. The function F_x assigns a

---

[4] Ibid.

number $F_x(\psi)$ to every $\psi \in \Phi$. This function fulfills the relation $F_x(a\psi + b\phi) = \bar{a}F_x(\psi) + \bar{b}F_x(\phi)$ for all $\psi, \phi \in \Phi$ and for all complex numbers $a, b \in \mathbb{C}$, which is the defining property of an anti-linear functional. But for calculations here, generalized eigenvectors can be treated as if they were proper eigenvectors, and $\langle\psi|x\rangle$ can be treated as if it were a scalar product $\langle\psi|x\rangle$. The distinction between eigenvectors $|a_n\rangle$ and generalized eigenvectors $|x\rangle$ is that the sums over $|a_n\rangle$ in (A.4.23) must be replaced by integrals as in (A.4.24).

The generalized basis expansion (A.4.22a) does not hold for all vectors ψ for which the discrete basis vector expansion (A.4.20) is correct. Instead it is valid only for a subset of the Hilbert space which here is chosen as the Schwartz space Φ. This is due to the fact that (A.4.26) does not hold for all (Lebesgue square integrable) functions but only for "well-behaved" functions $\psi(x)$. An example of such functions are the Schwartz-space functions that are continuous, infinitely differentiable and have derivatives of any order that decrease faster than any inverse power of x.

Although the above discussion has focused on the eigenket $|x\rangle$, where x usually represents position, the statements apply equally well to the eigenkets $|E\rangle$ and $|p\rangle$ etc. in (A.4.21a) and (A.4.21b).

The components $\psi(x) = \langle x|\psi\rangle$, $\psi(p) = \langle p|\psi\rangle$, and $\psi(E) = \langle E|\psi\rangle$, which are also called wave functions, must satisfy certain conditions in quantum mechanics.

A.4.4 Working with Eigenvectors and Basis Vector Expansions

The eigenvector expansion

$$\psi = \sum_{n=1}^{\infty} |a_n\rangle\langle a_n|\psi\rangle \,, \tag{A.4.27}$$

associates with every vector ψ an infinite sequence of numbers $\{\langle a_n|\psi\rangle \,, n = 1, 2, 3\ldots\}$. The coordinates $c_n = \langle a_n|\psi\rangle$ are, in general, complex numbers and are a function of the discrete variable n. The vector $\psi = 0$ iff all of its coordinates are zero. Equivalently, the vectors ϕ and ψ are equal if all their coordinates are equal. The set of eigenvectors $|a_n\rangle$ in (A.4.27) is thus a complete system of eigenvectors. The set of eigenvalues $\{a_n, n = 1, 2, \ldots\}$ is called the spectrum of the operator A.

In analogy with (A.4.10), the square of the norm of a vector ψ is given by

$$\|\psi\|^2 \equiv (\psi, \psi) = \sum_{n=1}^{\infty} |c_n|^2 = \sum_{n=1}^{\infty} |\langle a_n|\psi\rangle|^2 = \sum_{n=1}^{\infty} \langle a_n|\psi\rangle^* \langle a_n|\psi\rangle \,. \tag{A.4.28}$$

If the vector has a finite norm (finite length) then

$$\sum_{n=1}^{\infty} |c_n|^2 = \sum_{n=1}^{\infty} |\langle a_n|\psi\rangle|^2 < \infty \,. \tag{A.4.29}$$

The space of square summable sequences, which is the space of all vectors ψ with components that fulfill (A.4.29), is the Hilbert space $\mathscr{H}$.

Using the fact that the identity operator $\mathbb{1}$ satisfies $\psi = \mathbb{1}\,\psi$ and $|a_n\rangle = \mathbb{1}\,|a_n\rangle$, it is possible to omit the vector ψ from both sides of (A.4.27) because the equation is true for any $\psi \in \Phi$ and write it as an equation for operators,

$$\mathbb{1} = \sum_{n=1}^{\infty} |a_n\rangle\langle a_n| \,. \tag{A.4.30}$$

Equation (A.4.30) is called the completeness relation for the basis system $\{|a_n\rangle\}$ or the spectral resolution of the identity operator. Using (A.4.30) the scalar product of two vectors $\phi, \psi \in \Phi$ can be expressed as an infinite sum,

$$(\phi, \psi) = (\phi, \mathbb{1}\,\psi) = \sum_{n=1}^{\infty} \langle \phi, |a_n\rangle\langle a_n|\psi\rangle, = \sum_{n=1}^{\infty} \langle \phi|a_n\rangle\langle a_n|\psi\rangle = \sum_{n=1}^{\infty} \langle a_n|\phi\rangle^* \langle a_n|\psi\rangle. \tag{A.4.31}$$

Equation (A.4.31) is the analogue of (A.4.9) in $\Re_3$. Here, since the space is complex, the scalar product is the sum of the products of the components of one vector with complex conjugate of the components of the other.

Equation (A.4.31) makes sense only if the sum converges. It is possible to show that if (A.4.29) is fulfilled for all vectors, in particular for the two vectors φ and ψ in (A.4.31) then

$$|(\varphi, \psi)| = |\sum_{n=1}^{\infty} (a_n, \varphi)^* \langle a_n|\psi\rangle| < \infty \,. \tag{A.4.32}$$

All vectors ψ, φ for which (A.4.32) is fulfilled have finite scalar products with each other and form the space $\mathscr{H}$.

The linear operators $A, B, C, \ldots,$ of (A.3.1b) and (A.3.2b) act in the space $\mathscr{H}$ and can be added and multiplied according to (A.3.3b)–(A.3.5b), forming an associative algebra. As is justified in the various chapters, the vectors $\varphi, \psi, \ldots, \in \mathscr{H}$ and the operators $A, B, \ldots, |\varphi\rangle\langle\varphi|, |\psi\rangle\langle\psi| \ldots$ represent quantum physical states and observables, respectively. Quantities such as $|(\varphi, \psi)|^2$, $|(\varphi, A\varphi)|$, $|(\varphi, A\psi)|$, and $|(\varphi, AB\psi)|$ are Born probabilities that describe the quantum physical quantities that are extracted from experiments. Since such quantities must be finite, it is necessary to require not only that (A.4.29) be finite, but also that $(\varphi, A^r\varphi)$, $(\varphi, B^s\varphi)$, $(\varphi, AB^r\varphi) \ldots, r, s = 1, 2, 3, \ldots$, have finite absolute values.

A space Φ that is "better" than a Hilbert space $\mathscr{H}$ is required because all operators $A, B, \ldots$ that represent observables for the quantum physical system under consideration and any arbitrary power r of the operators $A, B, \ldots$ must be well-defined in the space Φ. Thus the vector $A^r\psi$ must also have a finite norm. To determine the restriction this imposes, the square of the norm of $A^r\psi$, $B^s\psi$ is calculated.

Expressing the vector ψ in terms of the eigenvectors $|a_n\rangle$ of the operator A as given in (A.4.27) and then applying the operator A,

$$A\psi = A\sum_{n=1}^{\infty}|a_n\rangle\langle a_n|\psi\rangle = \sum_{n=1}^{\infty}a_n|a_n\rangle\langle a_n|\psi\rangle\,. \tag{A.4.33}$$

Since (A.4.33) is true for any $\psi \in \Phi$, ψ can be omitted in (A.4.33) just as was done in arriving at (A.4.30). Then (A.4.33) becomes the operator relation

$$A = \sum_{n=1}^{\infty}a_n|a_n\rangle\langle a_n| = \sum_{n=1}^{\infty}a_n\Lambda_n\,. \tag{A.4.34}$$

Thus a linear operator is the sum of projection operators $\Lambda_n = |a_n\rangle\langle a_n|$ multiplied by the respective eigenvalues a_n that are real if $A = A^\dagger$. Equation (A.4.34) is called the spectral resolution of the operator A.

An operator B that does not commute with A (i.e. for which $B\,A - A\,B \neq 0$) cannot be written in terms of the eigenvectors $|a_n\rangle$ of the operator A in the form (A.4.34) because $B\,|a_n\rangle \neq b_n|a_n\rangle$. However, if B is self-adjoint and has a discrete spectrum, then it can be expressed in terms of its eigenvectors $|b_n\rangle$ that satisfy $B\,|b_n\rangle = b_n|b_n\rangle$. Results analogous to (A.4.30) and (A.4.34) are then immediately obtained:

$$\mathbb{1} = \sum_{n=1}^{\infty}|b_n\rangle\langle b_n|, \qquad B = \sum_{n=1}^{\infty}b_n|b_n\rangle\langle b_n| \tag{A.4.35}$$

Using the eigenvectors of B as a basis system, every $\psi \in \Phi$ can as well be written as

$$\psi = \sum_{n=1}^{\infty}|b_n\rangle\langle b_n|\psi\rangle \tag{A.4.36}$$

which, of course, is just (A.4.27) in a different basis system $\{|b_n\rangle, n = 1, 2, \ldots\}$. In general the $|b_i\rangle$ and $|a_i\rangle$ are completely different vectors. Replacing ψ by $|a_i\rangle$ in (A.4.36), each basis vector $|a_i\rangle$ can be expressed as an infinite sum of basis vectors $|b_n\rangle$,

$$|a_i\rangle = \sum_{n=1}^{\infty}|b_n\rangle\langle b_n|a_i\rangle. \tag{A.4.37}$$

Using (A.4.33) it is possible to calculates $A^r\psi$ and $B^s\psi$ for any ψ and for $r, s = 1, 2, 3, \ldots$. Using the fact that A and, as a consequence A^r, is self-adjoint,

$$\|A^r\psi\|^2 = (A^r\psi, A^r\psi) = (\psi, A^{2r}\psi)\,. \tag{A.4.38}$$

Expressing ψ as an infinite sum of eigenstates of the operator A as given in (A.4.27),

$$\|A^r\psi\|^2 = \sum_{n=1}^{\infty}(\psi, A^{2r}|a_n\rangle\langle a_n|\psi\rangle, = \sum_{n=1}^{\infty} a_n^{2r}\langle\psi|a_n\rangle\langle a_n|\psi\rangle, \tag{A.4.39}$$

Thus the vector $A^r\psi$ is defined if the sum

$$\sum_{n=1}^{\infty} a_n^{2r}|\langle\psi|a_n\rangle|^2 < \infty \quad \text{for every } r = 1, 2, \ldots\,. \tag{A.4.40}$$

The norm of the vector $B^s\psi$ is calculated similarly.

From the preceding discussion it follows that the space of vectors on which all powers of operators A^r, B^s, ... are defined is the space of vectors with components $|\langle\psi|a_n\rangle|$, $|\langle\varphi|b_n\rangle|$ that are rapidly decreasing. These are the vectors of the space $\mathscr{H}$ for which not only (A.4.29) holds, but also for which the more stringent condition (A.4.40) is fulfilled for all operators A, B, ..., thereby ensuring that matrix elements of the form $(\varphi, A^r B^s\psi)$ are also defined.

This smaller space

$$\Phi \subset \mathscr{H} \tag{A.4.41}$$

is by hypothesis the space of states and of observables of quantum physical systems. These spaces can be defined mathematically such that it is possible to prove Dirac's continuous basis vector expansion (A.4.22a) as a mathematical theorem, the "Nuclear Spectral Theorem." The kets $|a\rangle, |x\rangle, |E\rangle, \ldots$ are not vectors in Φ or in $\mathscr{H}$, they are instead continuous, anti-linear functionals on the space Φ. The space of continuous, anti-linear functionals, which is the space containing these kets, is denoted by $\Phi^\times$. That is, the space of all anti-linear, continuous functionals on the space Φ (with respect to the convergence in Φ) are denoted by $\Phi^\times$. Since the space of continuous, anti-linear functionals $\mathscr{H}^\times$ on the Hilbert space $\mathscr{H}$ is again a Hilbert space, $\mathscr{H}^\times = \mathscr{H}$. From (A.4.41) the following triplet of spaces is obtained:

$$\Phi \subset \mathscr{H} \subset \Phi^\times\,. \tag{A.4.42}$$

This is called the Gelfand triplet or Rigged Hilbert Space. The vectors $\psi, \phi \in \Phi$ are called "well-behaved," and the Dirac kets $|x\rangle, |p\rangle, |E\rangle, \ldots$ are elements of $\Phi^\times$.

For practical calculations in physics, the underlying mathematics is not so important. But it is important to know that these mathematical objects are rigorously defined and to know their limitations and properties. The most important property of the Rigged Hilbert Space is the Nuclear Spectral Theorem that justifies the basis vector expansion (A.4.22) as Dirac foresaw.

The continuous analogue of (A.4.34), the spectral resolution of the operator Q, can be obtained by first operating on both sides of (A.4.22) with Q,

$$Q\psi = \int \mathrm{d}x\, Q|x\rangle\langle x|\psi\rangle \int \mathrm{d}x\, x|x\rangle\langle x|\psi\rangle\,,$$

and then omitting the arbitrary vector $\psi \in \Phi$:

$$Q = \int \mathrm{d}x\, x|x\rangle\langle x|\,. \tag{A.4.43}$$

Because (A.4.22) is true for any $\psi \in \Phi$, it is possible to omit ψ from the equation,

$$\mathbb{1} = \int \mathrm{d}x|x\rangle\langle x|\,. \tag{A.4.44}$$

Equation (A.4.44) is the continuous analogue of the discrete relation (A.4.30) and is called the completeness relation of the generalized basis system $\{|x\rangle\}$.

The scalar product of two elements $\varphi, \psi \in \Phi$ is then obtained from (A.4.44),

$$(\varphi, \psi) \equiv \langle\varphi|\psi\rangle = \langle\varphi|\mathbb{1}\,\psi\rangle = \int \mathrm{d}x\,\langle\varphi|x\rangle\langle x|\psi\rangle. \tag{A.4.45}$$

In (A.4.45)

$$\langle\varphi|x\rangle \equiv \langle\varphi, |x\rangle) = (|x\rangle, \varphi)^* \equiv \langle x|\varphi\rangle^* \tag{A.4.46}$$

is the (generalized) scalar product of $\varphi \in \Phi$ with the generalized basis vector $|x\rangle$. Using the standard notation $\psi(x) = \langle x|\psi\rangle$ and $\varphi^*(x) = \langle\varphi|x\rangle$, (A.4.45) can be rewritten in the form

$$(\varphi, \psi) = \int \mathrm{d}x\, \varphi^*(x)\psi(x), \tag{A.4.47}$$

the familiar form of the scalar product in function spaces (A.2.12).

Just as there are conditions on the components $c_n = \langle a_n|\psi\rangle$ for the discrete case, there are also corresponding conditions on the components $\psi(x) = \langle x|\psi\rangle$ for the continuous case. From (A.4.45), it immediately follows that the square of the norm of ψ is given by

$$\|\psi\|^2 = (\psi, \psi) = \int \mathrm{d}x\,\langle\psi|x\rangle\langle x|\psi\rangle = \int \mathrm{d}x\,|\psi(x)|^2. \tag{A.4.48}$$

The components $\psi(x)$ must therefore be square integrable functions for the norm to be finite. Furthermore, if the operator Q and an arbitrary power r of the operator Q are to be well-defined in the space Φ, the vector $Q^r\psi$ must also have a finite norm. Performing a calculation analogous to (A.4.39),

$$\|Q^r\psi\|^2 = (Q^r\psi, Q^r\psi\rangle = \int \mathrm{d}x\, x^{2r}|\psi(x)|^2. \tag{A.4.49}$$

If the norm $\|Q^r\psi\|$ is to be finite, from (A.4.39) it follows that $|\psi(x)|^2$ must decrease faster than any power of x.

Using (A.4.21c) the matrix elements of the self-adjoint operator Q with eigenkets $|x\rangle$ are

$$\langle x|Q|\psi\rangle = (Q|x\rangle, |\psi\rangle) = (x|x\rangle, |\psi\rangle) = x\langle x|\psi\rangle \quad \text{for all } \psi \in \Phi. \tag{A.4.50}$$

An operator P is now sought with matrix elements $\langle x|P|\psi\rangle$ between the continuous basis vector $|x\rangle$ and any $\psi \in \Phi$ that are given by

$$\langle x|P|\psi\rangle = \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}\langle x|\psi\rangle = \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}\psi(x). \tag{A.4.51}$$

Since an arbitrary power Q^rP^s (or P^sQ^r) is to be a well-defined operator, their matrix elements must be finite. Thus the components $\langle x|\varphi\rangle$ and $\langle x|\psi\rangle$ must fulfill the condition that

$$\langle\varphi|Q^rP^s|\psi\rangle = \int \mathrm{d}x\,\langle\varphi|Q^r|x\rangle\langle x|P^s|\psi\rangle \tag{A.4.52}$$

exists. Using (A.4.50) and (A.4.51)

$$|\langle\varphi|Q^rP^s|\psi\rangle| = |\int \mathrm{d}x\varphi^*(x)\,x^r\frac{1}{i^s}\frac{\mathrm{d}^s}{\mathrm{d}x^s}\psi(x)| < \infty \tag{A.4.53}$$

for all $r = 1, 2, \dots\ s = 1, 2, \dots$ and all φ, ψ.

Equation (A.4.53) reveals that the components in the $|x\rangle$-basis, called the position wave functions $\varphi(x)$ and $\psi(x)$, must fulfill the following condition: The products of the position wave functions and all their s derivatives must decrease faster than any power r of x. The infinitely differentiable, rapidly decreasing, smooth functions that fulfill these conditions are call Schwartz-space functions, and the space of these functions is called the Schwartz space.

It is possible to calculate the commutator of the operators Q and P defined by (A.4.50), (A.4.51) in the Schwartz space:

$$\begin{aligned}\langle x|[Q,P]|\psi\rangle &= \langle x|QP-PQ|\psi\rangle\\ &= x\frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}\langle x|\psi\rangle - \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}\langle x|Q|\psi\rangle = -\frac{1}{i}\langle x|\psi\rangle = i\langle x|\mathbb{1}|\psi\rangle. \qquad \text{(A.4.54)}\end{aligned}$$

The above formula is valid for every function $\langle x|\psi\rangle$ in Schwartz space, which means it is valid for every $|x\rangle$ and for every vector $\psi \in \Phi$. Thus it is true as an operator equation in Φ,

$$[Q,P] = i\mathbb{1}. \qquad \text{(A.4.55)}$$

The correspondence in the following table,

Space Φ		Schwartz Function Space of x
ψ	corresponds to	$\langle x\|\psi\rangle$
Q	corresponds to	operator that multiplies by x
P	corresponds to	differentiation operator $\frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}$

is called a realization of the space Φ.

It is important to emphasize that no proofs have been given in the above discussion . Formalism has been presented based on (A.4.27) and (A.4.43), which have been written in analogy to the basis vector expansion in $\Re_3$. But the statements (A.4.27) and (A.4.43), which are special cases of the Nuclear Spectral Theorem, are far from trivial and require proofs. In fact the Nuclear Spectral Theorem is one of the more important mathematical theorems, with much of this section being a consequence of it. But long before (A.4.43) was proved or even precisely formulated in terms of well-defined mathematical quantities, it was used successfully by Dirac in his formulation of quantum mechanics.

A.5 Realizations by Matrices and Functions

"Realizations" of linear, scalar-product spaces and linear operators are now briefly discussed. To illustrate what is meant by a "realization" it is convenient to return to the three-dimensional space $\Re_3$. There a vector $\mathbf{x}$ can be described by giving its magnitude and direction. Alternatively, it can be specified by its components

$x_i = \mathbf{e}_i \cdot \mathbf{V}$. Thus there are two alternate but equivalent descriptions which are written symbolically as

$$\mathbf{x} \leftrightarrow x_i = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{pmatrix}. \tag{A.5.1}$$

The coordinates or components, of course, depend on the chosen basis system.

In the same way, instead of using the vector $\psi \in \Phi$, it is possible to use the components $\langle n|\psi\rangle$ with respect to a discrete basis, which can be written as a column matrix.

$$\psi \leftrightarrow \langle n|\psi\rangle = \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \langle 3|\psi\rangle \\ \vdots \end{pmatrix} \tag{A.5.2}$$

Just as a vector in $\Re_3$ can be expressed in terms of coordinates or components with respect to different basis systems, it is possible to express ψ in terms of components with respect to a different basis. Instead of the basis $|n\rangle$, which are, say, eigenstates of the energy operator H, it is possible to use as a basis eigenvectors $|a_n\rangle$ of the operator A satisfying $A|a_n\rangle = a_n|a_n\rangle$. In terms of $|a_n\rangle$, the same vector is given by an entirely different column matrix.

$$\psi \leftrightarrow \langle a_n|\psi\rangle = \begin{pmatrix} \langle a_1|\psi\rangle \\ \langle a_2|\psi\rangle \\ \langle a_3|\psi\rangle \\ \vdots \end{pmatrix}. \tag{A.5.3}$$

The column matrices $\langle n|\psi\rangle$ and $\langle a_n|\psi\rangle$ are related by an infinite-dimensional transformation matrix, which is usually so complicated that it is of no practical value. (See Problem A.14.) If the infinite column matrix appears more real to someone than the abstract vector ψ, then that person will speak of a realization of ψ by a column matrix and a realization of the space Φ by the space of column matrices.

When vectors are realized by column matrices, operators are realized by quadratic matrices. To illustrate this concept, the action of an arbitrary operator B on the vector ψ is calculated using the expansion (A.4.13),

$$B\psi = \sum_{n=1}^{\infty} B|n\rangle\langle n|\psi\rangle. \tag{A.5.4}$$

Taking the scalar product of the above equation with the basis vector $|m\rangle$ yields

$$\langle m|B\psi\rangle = \sum_{n=1}^{\infty} \langle m\ |B|n\rangle\langle n|\psi\rangle. \tag{A.5.5}$$

The above relation can be written in matrix notation as follows:

$$\begin{pmatrix} \langle 1|B\psi\rangle \\ \langle 2|B\psi\rangle \\ \langle 3|B\psi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle 1|B|1\rangle & \langle 1|B|2\rangle & \langle 1|B|3\rangle & \ldots \\ \langle 2|B|1\rangle & \langle 2|B|2\rangle & \langle 2|B|3\rangle & \ldots \\ \langle 3|B|1\rangle & \langle 3|B|2\rangle & \langle 3|B|3\rangle & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \langle 3|\psi\rangle \\ \vdots \end{pmatrix} \tag{A.5.6}$$

The numbers $\langle m|B|n\rangle$ form an infinite-dimensional quadratic matrix which is called the matrix of the operator B with respect to the basis $|n\rangle$. An orthonormal basis is almost always used in calculating the matrix of an operator.

For any two vectors $\psi, \varphi\ \in\ \Phi$ and an operator B, the scalar product of $B\psi$ with φ, namely $\langle\varphi|B\psi\rangle$, plays an important role in physics and is called the matrix element of the operator B between the vectors φ and ψ.

The vector $\psi\ \in\ \Phi$ can be expanded in terms of a continuous basis instead of a discrete basis. Then instead of the correspondence (A.5.2),

$$\psi\ \leftrightarrow\ \langle x|\psi\rangle. \tag{A.5.7}$$

The column matrix $\langle x|\psi\rangle$ has continuously infinite rows with one row for each value of x. With the association (A.5.7), the space Φ is realized by the space of "well-behaved" functions.

A.6 Summary

Quantum mechanics can be expressed in terms of differential operators using the Schrödinger equation. Equivalently, it can be expressed in terms of matrices using matrix mechanics. These two formulations of quantum mechanics are not distinct theories but are merely two different representations of quantum mechanics obtained from the general formulation by taking matrix elements with respect to different basis systems. In its most general form, quantum mechanics is formulated in terms of linear operators on a linear, scalar-product space.

A linear space Φ possesses the following ten properties where φ, ψ, and χ are elements or vectors in Φ; $0\ \in\ \Phi$ is the null element; and a, b are complex numbers:

1. $\varphi + \psi = \psi + \varphi$
2. Addition is associative: $(\varphi + \psi) + \chi = \varphi + (\psi + \chi)$

3. The null element satisfies $0 + \psi = \psi$
4. $b(\psi) = b\psi \in \Phi$
5. $a(b\psi) = (ab)\psi$
6. $\mathbb{1}\psi = \psi$
7. $0\psi = 0$
8. $b(\varphi + \psi) = b\varphi + b\psi$
9. $(a + b)\psi = a\psi + b\psi$
10. $-\mathbb{1}\psi = -\psi$

A linear, scalar-product space Φ possesses the ten properties that characterize a linear space plus the following four properties that define a scalar-product (ψ, φ) on a linear space:

1. $(\psi, \psi) \geq 0,\ (\psi, \psi) = 0$ if $\psi = 0$
2. $(\psi, \varphi) = (\varphi, \psi)^*$
3. $a(\psi, \varphi) = (\psi, a\varphi)$
4. $(\psi + \varphi, \chi) = (\psi, \chi) + (\varphi, \chi)$

Linear operators A and B on a linear, scalar-product space Φ possess the following nine properties:

1. $A(\psi) \equiv A\psi \in \Phi$
2. $A(\psi + \varphi) = A\psi + A\varphi$
3. $A(a\psi) = a(A\psi)$
4. $(A + B)\psi = A\psi + B\psi$
5. $(aA)\psi = a(A\psi)$
6. $(AB)\psi = A(B\psi)$
7. The zero operator 0 satisfies $0\psi = 0 \in \Phi$.
8. The identity operator $\mathbb{1}$ satisfies $\mathbb{1}\psi = \psi$.
9. The adjoint of the operator A, denoted $A^\dagger$, is defined by $(\psi, A\varphi) = (A^\dagger\psi, \varphi)$.

The basis vector expansion for quantum theory has been motivated by starting with the simplest case and then generalizing. First a vector in three-dimensional space is expanded in terms of its components,

$$\mathbf{x} = \sum_{i=1}^{3} \mathbf{e}_i x^i = \sum_{i=1}^{3} \mathbf{e}_i (\mathbf{e}_i | \mathbf{x}\rangle,$$

where x^i is real. Just as vectors in three-dimensional space can be expressed as components along various sets of three linearly independent axes, vectors in a linear, scalar-product space can also be expressed as components along various sets of linearly independent axes or vectors. In a complex, N-dimensional, linear, scalar-product space the above equation is generalized to

$$\psi = \sum_{i=1}^{N} |e_i\rangle c_i = \sum_{i=1}^{N} |e_i\rangle\langle e_i|\psi\rangle\,,$$

where c_i is complex. Finally, the above equation is generalized to infinite dimensions, $N \to \infty$,

$$\psi = \sum_{n=1}^{\infty} |e_n\rangle c_n = \sum_{n=1}^{\infty} |e_n\rangle\langle e_n|\psi\rangle , \tag{A.6.1}$$

where the c_n are complex. The set of all vectors with components c_n that are square summable,

$$\sum_{n=1}^{\infty} |c_n|^2 < \infty ,$$

is called the Hilbert space $\mathscr{H}$.

For finite-dimensional spaces, a set of basis functions $\{|e_i\rangle\}$ can be formed from the eigenvectors of any self-adjacent operator A:

$$|e_i\rangle = |a_i\rangle \qquad \text{where} \qquad A|a_i\rangle = a_i|a_i\rangle.$$

For the infinite-dimensional spaces there are some operators that do not have a discrete set of eigenvectors (called discrete spectrum). Then there exists a continuous set of eigenvectors satisfying

$$A|a\rangle, = a|a\rangle, \qquad M \leq a \leq N.$$

The vector ψ can be expanded in terms of $|a\rangle$,

$$\psi = \int \mathrm{d}a|a\rangle\langle a|\psi\rangle,$$

which is justified mathematically by the Nuclear Spectral Theorem and is the continuous generalization of (A.6.1).

Problems

For Sect. A.2

A.1 Consider the two functions $g_1(x)$ and $g_2(x)$ on the interval $-\infty \leq x < \infty$,

$$g_1(x) = A_1 e^{-|x|}, \qquad g_2(x) = A_2(a + x^2)e^{-|x|},$$

where the constants A_1 and A_2 are real.

(a) Determine the constant a such that $g_1(x)$ and $g_2(x)$ are orthogonal.
(b) Determine the constants A_1 and A_2 such that $g_1(x)$ and $g_2(x)$ are normalized.

A.2 Show that if the vectors $\psi, \varphi \in \Phi$ are represented by the respective column matrices

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_N \end{pmatrix}, \quad \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \vdots \\ \varphi_N \end{pmatrix},$$

where the ψ_i and ϕ_i are complex numbers, then the scalar product defined by

$$(\psi, \varphi) = \sum_{i=1}^{N} \psi_i^* \varphi_i$$

satisfies the rules (A.2.13c)–(A.2.16c).

For Sect. A.3

A.3 Show that rules (A.3.1b)–(A.3.5b) are satisfied if the vectors $\psi, \varphi \in \Phi$ are represented by column matrices as given in Problem A.2, and that an arbitrary linear operator A is represented by an $N \times N$ matrix A_{ij} such that the action of A on ψ is represented by

$$A\psi = \sum_{j=1}^{\infty} A_{ij} \psi_j \ .$$

The operator B is defined analogously.

A.4 If the vectors in a scalar-product space are represented by column matrices, and the operator A is represented by the matrix A_{ij} as given in Problems A.2 and A.3, by what is the operator $A^\dagger$ represented? Hint: Use a procedure similar to that employed in obtaining (A.3.6a).

A.5 Using (A.3.5b) and the definition (A.3.6b) of the adjoint of an operator, show that $(AB)^\dagger = B^\dagger A^\dagger$.

A.6 Let A and B be hermitian operators and c be an arbitrary, complex number. Under what conditions is each of the following operators hermitian?

(a) cA
(b) $cA + cB$
(c) cAB
(d) $c\{A, B\} = c(AB + BA)$
(e) $c[A, B] = c(AB - BA)$

A.7 Calculate the four possible matrix elements of the operator $-i\dfrac{\mathrm{d}}{\mathrm{d}x}$ between the functions

$$f_1(x) = \pi^{-1/4} e^{-x^2/2} \quad \text{and} \quad f_2(x) = \sqrt{2}\,\pi^{-1/4}\, x e^{-x^2/2},$$

where $-\infty < x < \infty$.

A.8 On the interval $-\infty < x < \infty$, consider the operator A, where

$$A = \frac{\mathrm{d}^2}{\mathrm{d}x^2} - x^2,$$

and the two functions

$$\psi_1(x) = e^{-x^2/2}, \qquad \psi_2(x) = x e^{x^2/2}.$$

(a) Show that $A = A^\dagger$.
(b) Show that

$$A\psi_i(x) = a_i \psi_i(x), \quad i = 1, 2 .$$

The constant a_i is called the eigenvalue of the operator A when it acts on $\psi_i(x)$, and $\psi_i(x)$ is called an eigenfunction of the operator A. What are the values a_1 and a_2?

(c) Using only the facts that $A^\dagger = A$ and $a_1 \neq a_2$, explain why the scalar product of $\psi_1(x)$ and $\psi_2(x)$ must be zero. Explicitly calculate the scalar product and verify that it is indeed zero.

For Sect. A.4

A.9 Let $\psi \in \Phi$ be normalized to unity. If $|n\rangle$ is a basis system of eigenvectors of an observable with a discrete spectrum, show that the components of ψ with respect to this basis fulfill the condition

$$\sum_n |\langle n|\psi\rangle|^2 = 1\ .$$

A.10 If $|e_i\rangle$ is a normalized eigenvector with eigenvalue a_i, show that

$$|e_n'\rangle = e^{i\omega}|e_n\rangle, \qquad \omega \in \Re\,,$$

is also a normalized eigenvector with the same eigenvalue a_i.

For Sect. A.5

A.11 Let A be a hermitian operator and $|n\rangle$ be a discrete basis system. Show that the matrix of A with respect to this basis system,

$$\begin{pmatrix} \langle 1|A|1\rangle & \langle 1|A|2\rangle & \dots \\ \langle 2|A|1\rangle & \langle 2|A|2\rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

is a hermitian matrix. That is, show that $\langle m|A|n\rangle = \langle n|A|m\rangle^*$.

A.12 A matrix T is said to be orthogonal if $T^{\mathrm{t}} = T^{-1}$. Here T^{t} is the transposed matrix, $T^{\mathrm{t}}_{mn} = T_{nm}$, and T^{-1} is the inverse matrix defined by

$$\sum_r T_{nr}(T^{-1})_{rm} = \delta_{nm}\ .$$

(a) Show that the matrix

$$T = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

is orthogonal.

(b) Calculate the determinant of T.
(c) Apply the matrix T to the vector

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

using the rules of matrix multiplication. Using a sketch, show that $T\mathbf{r}$ is the vector obtained by rotating the vector $\mathbf{r}$ around the z axis by an angle θ.

A.13 A matrix $\mathscr{P}$ is called a projection matrix if

$$\mathscr{P}^2 = \mathscr{P} .$$

(a) Show that the matrix

$$\mathscr{P}' = \begin{pmatrix} \sin^2\theta & -\sin\theta\cos\theta & 0 \\ -\sin\theta\cos\theta & \cos^2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is a projection matrix.
(b) Apply the projection matrix $\mathscr{P}'$ to the vector $\mathbf{r}' = T\mathbf{r}$ obtained in Problem A.11, and discuss the result.
(c) If the rotation T is applied to the vector $\mathbf{r}' = T\mathbf{r}$ calculated in Problem A.11, a simple result is found that suggests defining a new matrix

$$\mathscr{P} = T^{-1}\mathscr{P}'T .$$

Calculate $\mathscr{P}$ and verify that it is a projection matrix.
(d) Describe the geometrical meaning of the matrix $\mathscr{P}$. In particular, determine the subspace of the three-dimensional space on which $\mathscr{P}$ projects.
(e) Describe the geometrical meaning of the matrix $\mathscr{P}'$ and determine the subspace upon which it projects.

A.14 Let two basis systems of the linear, scalar-product space Φ be denoted by $|a_n\rangle$ and by $|b_\nu\rangle$. Show that the components of a vector $\varphi \in \Phi$ with respect to the basis system $|b_\nu\rangle$, the $(b_\nu|\phi\rangle$, can be obtained from the components $\langle a_n|\varphi\rangle$ with respect to the other basis system $|a_n\rangle$ by the matrix transformation

$$\langle b_\nu|\varphi\rangle = \sum_n \langle b_\nu|a_n\rangle\langle a_n|\varphi\rangle ,$$

where $\langle b_\nu|a_n\rangle$ is the scalar product of the basis vector $|b_\nu\rangle$ with the basis vector $|a_n\rangle$. To show this, use the fact that every vector $|a_n\rangle$ can be expanded with respect to the basis system $|b_\nu\rangle$.

Index

A. Bohm et al., *Quantum Physics*, https://doi.org/10.1007/978-94-024-1760-9